DOCUMENT RESUME

ED 099 184 SE 016 990

AUTHOR	Martin, Nancy
TITLE	Convergence Properties of a Class of Probabilistic Adaptive Schemes Called Sequential Reproductive Plans. Psychology and Education Series, Technical Report No. 210.
INSTITUTION	Stanford Univ., Calif. Inst. for Mathematical Studies in Social Science.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	31 Jul 73
NOTE	63p.
EDRS PRICE	MF-$0.75 HC-$3.15 PLUS POSTAGE
DESCRIPTORS	Computers; *Computer Science Education; Educational Research; *Genetics; Linear Programing; *Mathematical Models; Reproduction (Biology); Science Education; Social Sciences

ABSTRACT

Presented is a technical report concerning the use of a mathematical model describing certain aspects of the duplication and selection processes in natural genetic adaptation. This reproductive plan/model occurs in artificial genetics (the use of ideas from genetics to develop general problem solving techniques for computers). The reproductive plan is a sequential stochastic process involving n-tuples (corresponding to chromosomes in genetics) which may be simple numeric constants or complex structures such as computer algorithms. The plan also involves a sequence of probability distributions defined over n-tuples. The report consists of five chapters: introduction; reproductive plans; deterministic problem bases; a chapter divided into sections on the search for an arena, the linear additive model, the linear models and pure problem bases; and conclusions. An appendix illustrating the theorem involved and a list of references conclude the report. (PEB)

CONVERGENCE PROPERTIES OF A CLASS OF PROBABILISTIC ADAPTIVE SCHEMES CALLED SEQUENTIAL REPRODUCTIVE PLANS

BY

NANCY MARTIN

TECHNICAL REPORT NO. 210

JULY 31, 1973

PSYCHOLOGY AND EDUCATION SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

TECHNICAL REPORTS

PSYCHOLOGY SERIES

BEST COPY AVAILABLE

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

(Place of publication shown in parentheses; if published title is different from title of Technical Report, this is also shown in parentheses.)

(For reports no. 1 - 44, see Technical Report no. 125.)

50 R. C. Atkinson and R. C. Calfee. Mathematical learning theory. January 2, 1963. (In B. B. Wolman (Ed.), Scientific Psychology. New York: Basic Books, Inc., 1965. Pp. 254-275)

51 P. Suppes, E. Crothers, and R. Weir. Application of mathematical learning theory and linguistic analysis to vowel phoneme matching in Russian words. December 28, 1962.

52 R. C. Atkinson, R. Calfee, G. Sommer, W. Jeffrey and R. Shoemaker. A test of three models for stimulus compounding with children. January 29, 1963. (J. exp. Psychol., 1964, 67, 52-58)

53 E. Crothers. General Markov models for learning with inter-trial forgetting. April 8, 1963.

54 J. L. Myers and R. C. Atkinson. Choice behavior and reward structure. May 24, 1963. (Journal math. Psychol., 1964, 1, 170-203)

55 R. E. Robinson. A set-theoretical approach to empirical meaningfulness of measurement statements. June 10, 1963.

56 E. Crothers, R. Weir and P. Palmer. The role of transcription in the learning of the orthographic representations of Russian sounds. June 17, 1963.

57 P. Suppes. Problems of optimization in learning a list of simple items. July 22, 1963. (In Maynard W. Shelly, II and Glenn L. Bryan (Eds.), Human Judgments and Optimality. New York: Wiley. 1964. Pp. 116-126)

58 R. C. Atkinson and E. J. Crothers. Theoretical note: all-or-none learning and intertrial forgetting. July 24, 1963.

59 R. C. Calfee. Long-term behavior of rats under probabilistic reinforcement schedules. October 1, 1963.

60 R. C. Atkinson and E. J. Crothers. Tests of acquisition and retention, axioms for paired-associate learning. October 25, 1963. (A comparison of paired-associate learning models having different acquisition and retention axioms, J. math. Psychol., 1964, 1, 285-315)

61 W. J. McGill and J. Gibbon. The general-gamma distribution and reaction times. November 20, 1963. (J. math. Psychol., 1965, 2, 1-18)

62 M. F. Norman. Incremental learning on random trials. December 9, 1963. (J. math. Psychol., 1964, 1, 336-351)

63 P. Suppes. The development of mathematical concepts in children. February 25, 1964. (On the behavioral foundations of mathematical concepts. Monographs of the Society for Research in Child Development, 1965, 30, 60-96)

64 P. Suppes. Mathematical concept formation in children. April 10, 1964. (Amer. Psychologist, 1966, 21, 139-150)

65 R. C. Calfee, R. C. Atkinson, and T. Shelton, Jr. Mathematical models for verbal learning. August 21, 1964. (In N. Wiener and J. P. Schoda (Eds.), Cybernetics of the Nervous System: Progress in Brain Research. Amsterdam, The Netherlands: Elsevier Publishing Co., 1965. Pp. 333-349)

66 L. Keller, M. Cole, C. J. Burke, and W. K. Estes. Paired associate learning with differential rewards. August 20, 1964. (Reward and information values of trial outcomes in paired associate learning. (Psychol. Monogr., 1965, 79, 1-21)

67 M. F. Norman. A probabilistic model for free-responding. December 14, 1964.

68 W. K. Estes and H. A. Taylor. Visual detection in relation to display size and redundancy of critical elements. January 25, 1965, Revised 7-1-65. (Perception and Psychophysics, 1966, 1, 9-16)

69 P. Suppes and J. Donio. Foundations of stimulus-sampling theory for continuous-time processes. February 9, 1965. (J. math. Psychol., 1967, 4, 202-225)

70 R. C. Atkinson and R. A. Kinchla. A learning model for forced-choice detection experiments. February 10, 1965. (Br. J. math stat. Psychol., 1965, 18, 184-206)

71 E. J. Crothers. Presentation orders for items from different categories. March 10, 1965.

72 P. Suppes, G. Groen, and M. Schlag-Rey. Some models for response latency in paired-associates learning. May 5, 1965. (J. math. Psychol., 1966, 3, 99-128)

73 M. V. Levine. The generalization function in the probability learning experiment. June 3, 1965.

74 D. Hansen and T. S. Rodgers. An exploration of psycholinguistic units in initial reading. July 6, 1965.

75 B. C. Arnold. A correlated urn-scheme for a continuum of responses. July 20, 1965.

76 C. Izawa and W. K. Estes. Reinforcement-test sequences in paired-associate learning. August 1, 1965. (Psychol. Reports, 1966, 18, 879-919)

77 S. L. Blehart. Pattern discrimination learning with Rhesus monkeys. September 1, 1965. (Psychol. Reports, 1966, 19, 311-324)

78 J. L. Phillips and R. C. Atkinson. The effects of display size on short-term memory. August 31, 1965.

79 R. C. Atkinson and R. M. Shiffrin. Mathematical models for memory and learning. September 20, 1965.

80 P. Suppes. The psychological foundations of mathematics. October 25, 1965. (Colloques Internationaux du Centre National de la Recherche Scientifique. Editions du Centre National de la Recherche Scientifique. Paris: 1967. Pp. 213-242)

81 P. Suppes. Computer-assisted instruction in the schools: potentialities, problems, prospects. October 29, 1965.

82 R. A. Kinchla, J. Townsend, J. Yellott, Jr., and R. C. Atkinson. Influence of correlated visual cues on auditory signal detection. November 2, 1965. (Perception and Psychophysics, 1966, 1, 67-73)

83 P. Suppes, M. Jerman, and G. Groen. Arithmetic drills and review on a computer-based teletype. November 5, 1965. (Arithmetic Teacher, April 1966, 303-309.

84 P. Suppes and L. Hyman. Concept learning with non-verbal geometrical stimuli. November 15, 1968.

85 P. Holland. A variation on the minimum chi-square test. (J. math. Psychol., 1967, 3, 377-413).

86 P. Suppes. Accelerated program in elementary-school mathematics -- the second year. November 22, 1965. (Psychology in the Schools, 1966, 3, 294-307)

87 P. Lorenzen and F. Binford. Logic as a dialogical game. November 29, 1965.

88 L. Keller, W. J. Thomson, J. R. Tweedy, and R. C. Atkinson. The effects of reinforcement interval on the acquisition of paired-associate responses. December 10, 1965. (J. exp. Psychol., 1967, 73, 268-277)

89 J. I. Yellott, Jr. Some effects on noncontingent success in human probability learning. December 15, 1965.

90 P. Suppes and G. Groen. Some counting models for first-grade performance data on simple addition facts. January 14, 1966. (In J. M. Scandura (Ed.), Research in Mathematics Education. Washington, D. C.: NCTM, 1967. Pp. 35-43.

91 P. Suppes. Information processing and choice behavior. January 31, 1966.

92 G. Groen and R. C. Atkinson. Models for optimizing the learning process. February 11, 1966. (Psychol. Bulletin, 1966, 66, 309-320)

93 R. C. Atkinson and D. Hansen. Computer-assisted instruction in initial reading: Stanford project. March 17, 1966. (Reading Research Quarterly, 1966, 2, 5-25)

94 P. Suppes. Probabilistic inference and the concept of total evidence. March 23, 1966. (In J. Hintikka and P. Suppes (Eds.), Aspects of Inductive Logic. Amsterdam: North-Holland Publishing Co., 1966. Pp. 49-65.

95 P. Suppes. The axiomatic method in high-school mathematics. April 12, 1966. (The Role of Axiomatics and Problem Solving in Mathematics. The Conference Board of the Mathematical Sciences, Washington, D. C. Ginn and Co., 1966. Pp. 69-76.

(Continued on inside back cover)

CONVERGENCE PROPERTIES OF A CLASS OF PROBABILISTIC ADAPTIVE SCHEMES CALLED SEQUENTIAL REPRODUCTIVE PLANS

by

Nancy Martin

TECHNICAL REPORT NO. 210

July 31, 1973

PSYCHOLOGY AND EDUCATION SERIES

This research was supported by National Science Foundation Grant No. GJ-443X to the Institute for Mathematical Studies in the Social Sciences, Stanford University, and by National Science Foundation Grant No. GJ-29989X to the Logic of Computers Group, University of Michigan.

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

ABSTRACT

CONVERGENCE PROPERTIES OF A CLASS OF PROBABILISTIC ADAPTIVE SCHEMES CALLED SEQUENTIAL REPRODUCTIVE PLANS

by

Nancy Martin

Chairman: John H. Holland

A reproductive plan is a mathematical model describing certain aspects of the duplication and selection processes in natural genetic adaptation. These models occur in artificial genetics, which is the use of ideas from genetics to develop general problem solving techniques for computers.

A reproductive plan is a sequential stochastic process involving n-tuples which correspond to chromosomes in genetics. The individual elements of the n-tuples, which correspond to genes, may be simple numeric constants or may be such complex structures as computer algorithms. The plan also involves a sequence of probability distributions defined over the n-tuples.

At each step of the stochastic process, one of the n-tuples is selected using the current probability distribution. The "value" of the selected n-tuple is then obtained from an external function or subroutine. This value is then used to define a new probability distribution for the next step of the process.

A particular reproductive plan is said to converge if the distributions developed at each step converge to a distribution which selects the most valuable n-tuple. We analyse the convergence properties of several subclasses of reproductive plans. We show that in a suitably restricted problem domain one subclass, SR1 plans, converges. We also show that the convergence is not fast enough to achieve finite loss.

By relating reproductive plans to a class of models used in mathematical psychology, linear additive models, we show that several subclasses of reproductive plans do not converge.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to John Holland for introducing me to the area of genetic adaptation, for many helpful and stimulating conversations on this work and for his encouragement, support and patience.

Several times when I have been ready to leave academic work, Patrick Suppes not only encouraged me to continue but helped with financial assistance and exciting projects. He has been a great source of inspiration for many years and I will always be grateful.

I would like to thank Joyce Friedman for her careful reading of this manuscript and her many useful suggestions. I am extremely grateful for the support and encouragement she has given me.

I would also like to thank the other members of my reading committee, Bernard Zeigler, Robert Bjork and James Greeno for their helpful comments and suggestions.

There are many other people who have contributed to the completion of this work. I would particularly like to thank Tom Cover, Mary Harris, and Leslie Hefner. Jan McDougall and Monna Whipp have done an excellent job typing the manuscript.

Cleve Moler has contributed to my education and to the completion of this manuscript more than any other individual. I am grateful for the many happy hours we have spent studying and working together. His many suggestions have helped to make this manuscript more readable.

To Cleve, Teresa, and Kathryn I am greatly indebted for their patience, understanding and love.

This work was supported in part by National Science Foundation Grant No. GJ-443X to the Institute for Mathematical Studies in the Social

Sciences, Stanford University, and by National Science Foundation Grant No. GJ-29989X to the Logic of Computers Group, University of Michigan.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

The Logic of Computers Group at The University of Michigan has been studying methods of applying the techniques of natural genetic adaptation to develop adaptive techniques for problem solving with computers. We will use the term *artificial genetics* to refer to this process. We give here a very brief description of the genetic approach to problem solving in order to demonstrate the origins of the adaptive procedures we have investigated. For a thorough introduction we recommend Holland [to be published]. Chapter 1 of Hollstien (1971) relates the artificial genetic approach to adaptive control processes. Chapter 1 of Cavicchio (1970) relates the artificial genetic approach to pattern recognition and problems in artificial intelligence.

We separate the world of the adaptive process into two parts: the adaptive algorithm and the environment with which it must interact. If an algorithm is to be adaptive then something internal to the structure of the algorithm must change as time progresses. The first problem is to find an adequate method of representing that which will change. As in natural genetics, artificial genetics assumes that there is a set, $\mathcal{A}$, of "chromosomes". Each chromosome or string $A \in \mathcal{A}$ is an n-tuple, $(a_1,\ldots,a_n)$ where the a_i can be simple constants or quite complex structures such as the instructions for a computer subroutine. The a_i are referred to as genes. Hollstien (1971) has investigated the use of Gray codes (also known as reflected codes) and Hash codes to make the representation of information in the n-tuple more efficient.

We may view the problem of adaptation as a transformation problem: from a set of strings $\mathcal{A}_t$ obtain a new set of strings $\mathcal{A}_{t+1}$ by applying a set of operators based on some evaluation of the strings in $\mathcal{A}_t$. We assume that external to the adaptive process there is an evaluation function that evaluates the elements of $\mathcal{A}_t$ and that the results of this evaluation are available to the genetic process. In an artificial genetic adaptive scheme the operators are modeled after the natural genetic operators: duplication, crossover, inversion, mutation and dominance.

In the approach taken by Holland (to be published), the process is divided into two stages. The duplication operator is applied to $\mathcal{A}_t$ to provide a new set $\mathcal{A}'_t$ containing multiple copies of some of the strings of $\mathcal{A}_t$. The number of copies made of a particular string depends on the evaluation of that string. So duplication is a copying process that does not alter the individual strings. In the second state of the process, the other operators are applied to the set $\mathcal{A}'_t$ to form a new set of strings. Then the evaluation process is used to reduce the size of this set to form $\mathcal{A}_{t+1}$.

The crossover operator is a function which takes two strings as arguments and creates two new strings by interchanging genes of one string with the corresponding genes of the other string. For example if

$$A = (a_1, \ldots, a_n) \quad \text{and}$$

$$B = (b_1, \ldots, b_n)$$

then the result of a simple crossover operator might be the strings

$$(a_1, \ldots, a_i, b_{i+1}, \ldots, b_n) \quad \text{and}$$

$$(b_1, \ldots, b_i, a_{i+1}, \ldots, a_n).$$

The result of a "double" crossover operator might be

$$(a_1,\ldots,a_i,b_{i+1},\ldots,b_k,a_{k+1},\ldots,a_n) \quad \text{and}$$

$$(b_1,\ldots,b_i,a_{i+1},\ldots,a_k,b_{k+1},\ldots,b_n).$$

The inversion operator is a function of a single string which reverses the order of some segment of the string. For example, the result of applying an inversion operator to string A above might be the string $(a_1,\ldots,a_i,a_k,a_{k-1},\ldots,a_{i+1},a_{k+1},\ldots,a_n)$. The mutation operator makes random changes in the genes. The dominance operator is only applicable when the chromosomes have a more specific structure. Essentially, it chooses which copy of a gene will be effective if there is more than one representation for the same gene in the string.

The mutation, crossover and inversion operators can also be applied at the gene level to change the structure of the individual a_i's. A review of some of the algebraic aspects of these operators is presented in Foo and Bosworth (1972).

Experimental work with artificial genetic adaptation has been carried out by Hollstien (1971), Cavicchio (1970), Bosworth, Foo and Zeigler (1972) and Dan Frantz (1972). Holland began the theoretical investigation by analyzing the duplication phase of the process in (1969) and (1970). The present work is limited to this first phase of the adaptive scheme.

In order to capture the notion of duplication theoretically, a new algorithm called a Reproductive Plan was developed which did not use the other operators. It was intended that this algorithm could act as a driver program for the total genetic scheme.

A Reproductive Plan is a sequential stochastic process. If we view adaptation as a decision theory problem, then we must decide at

each time step which elements of $\mathscr{A}$ to choose. The problem is to find a suitable process to maximize the sum of all the outcomes or to find a process that eventually only chooses the "best" element of $\mathscr{A}$, according to some measure of "best". One method of making this decision is to have a probability vector over the space $\mathscr{A}$ which changes with experience. There are many methods, both linear and nonlinear for changing the values of the probability vector according to past performances. Many of these methods have been explored using different terminology by mathematical psychologists in searching for a model of behavior. While we are not interested in modeling any actual observed behavior, we are interested in taking advantage of the analysis that has been made and extending it to our particular requirements.

Shapiro and Narendra (1969) have compared the performance of several mathematical psychology models in the problem of function optimization with noise. Norman (1970) has done extensive work in analyzing a model which we show is very close to the reproductive plans of Holland.

Non-probabilistic methods of choosing among the elements of $\mathscr{A}$ have also been studied. A special problem of the type we are considering is called the n-armed bandit problem. This problem is generally stated as follows: we are given n coins with unknown probabilities $p_1,\ldots,p_n$ of heads. At each time step we are to choose which coin to toss. The objective is to find a sequential decision procedure that maximizes the limiting proportion of heads. Robbins (1952) was one of the first statisticians to examine the problem of sequential testing. He developed a successful rule for the case $n = 2$ which used the sample mean of the previous t tosses to choose the coin for the $(t+1)$st toss. The rule included a provision that prevented a coin from only being tested a

finite number of times. Later Robbins (1956) restated the problem allowing only a fixed number of previous tosses to be used in the decision making process. This problem is referred to as the bandit problem with finite memory. The exact interpretation of the term finite memory has been discussed in the literature and related to automata theory. For example see Cover (1969) and Hellman and Cover (1970, 1971).

In the present work we develop the conditions for convergence of reproductive plans and relate this to the models of mathematical psychology and statistical decision theory.

CHAPTER 2

REPRODUCTIVE PLANS

In this chapter we define a class of algorithms for adaptation similar to those developed by Holland (1970). This class of algorithms treats adaptation as a decision theory problem. There is a set of possible strategies and a probability vector over the strategies. There is also an evaluation function which measures the "worth" of a strategy at a particular time. A procedure sequentially chooses a strategy according to the probability vector. The procedure does not have direct access to the evaluation function but receives the resulting value of the function applied to the strategy choosen. This value is used to update the probability vector for the next choice. The object is to have the probability of the "best" strategy in S approach one as the number of trials increases. If a procedure is such that the probability of choosing any particular strategy in S goes to one with the number of trials, we say the procedure converges to that strategy. We now give definitions to make these notions precise.

Definition 2.1: $\langle A,\mu,S\rangle$ is a *problem basis* where

- A is a set of possible or admissible structures,
- μ is a function that assigns to each structure in A a random variable. We restrict the choice to the set of finite real random variables whose moment generating functions exist,
- S is a set of strategies for choosing structures from the space A. A strategy in S is a method for determining a trajectory or set of trajectories through the space A.

Example 2.2: In a zero-sum, two person game, let the structures of A be the pure strategies for player one, $|A| = M$, $T = \{(p_1,\ldots,p_M) | 0 \leqq p_i \leqq 1, \sum_{i=1}^{M} p_i = 1\}$, let $f:S \to T$ be 1-1, onto and for $s \in S$, s chooses structure a_i of A with probability p_i where p_i is the ith component of the probability vector $f(s)$. Note that in this example the word "strategy" is being used in two senses: first in its usual game theoretic sense and second to refer to elements of S which are "strategies for choosing strategies". For $a \in A$, the random variable $\mu(a)$ would be the "payoff" function of a. If we assume that player two always uses the same mixed strategy, then $\mu(a)$ is a simple random variable as in Definition 2.1 and $\langle A,\mu,S\rangle$ is a problem basis. However, if player two changes his strategy with time $\mu(a)$ no longer satisfies Definition 2.1.

Definition 2.3: A problem basis $\langle A,\mu,S\rangle$ is *deterministic* if for each $a \in A$, $\mu(a)$ is a finite constant (hence its variance is zero), and the strategies $s \in S$ are not probabilistic.

Example 2.4: If the structures in the set A are vectors in n-space, then any one of the standard function maximization algorithms of numerical analysis would be a nonprobabilistic strategy for determining a new structure of A given one or more "previous" structures and their function values $\mu(a_i)$. A collection of such algorithms could be the set S in a deterministic problem basis.

Definition 2.5: A strategy $s \in S$ is a *pure strategy* if s always selects the same element of A.

Definition 2.6: A problem basis $\langle A,\mu,S\rangle$ is *pure* if S is a set of pure strategies and the variances of the random variables $\mu(a)$, $a \in A$, are nonzero.

The problem of maximizing a real valued function defined on a finite set of points where there is random noise in the function evaluation can be considered a pure problem basis. Here $A = \{a_1,\ldots,a_n\}$ is a set of n points on the real line, $\mu(a_i)$, $i = 1,\ldots,n$, is uniformly distributed with mean ρ_i in the interval $[\rho_i-2,\rho_i+2]$ and S is the set of pure strategies such that there is one and only one strategy in S for each element of A. This problem was discussed by Shapiro and Narendra (1969).

The n-armed bandit problem is an example of a pure problem basis with A the set of n arms,

$$\mu(a_i) = \begin{cases} 1 & \text{with probability } \rho_i \\ 0 & \text{with probability } 1-\rho_i. \end{cases}$$

Again, S is the set of pure strategies such that there is one and only one strategy in S for each element of A.

Definition 2.7: A *sequential adaptive scheme*, SAS, over the problem basis $\langle A,\mu,S\rangle$ is an algorithm for choosing

1. An initial starting structure of A denoted A_0
2. At each time step, t, a strategy from S to be used to obtain a new structure.

Definition 2.8: A *probabilistic sequential adaptive scheme*, PSAS, is an SAS with the strategy at time t chosen according to a probability distribution over S.

Definition 2.9: A PSAS over a problem basis $\langle A,\mu,S\rangle$ is said to *converge* to the set $H \subset S$ if $\lim_{t\to\infty} \sum_{s\in H} P_{s,t} = 1$ where $P_{s,t}$ is the probability of choosing strategy s at time t.

We will use the following notation:

t — time variable.

$N(t)$ — the number of distinct strategies an SAS has selected prior to time t.

s_j — the jth distinct strategy selected by an SAS.

$S(t)$ — the set of distinct strategies selected prior to time t. $S(t) = \{s_1, s_2, \ldots, s_{N(t)}\}$.

t_j — the time strategy s_j is first selected.

$A_s(k)$ — the structure that would result if strategy s were being selected by an SAS for the kth time.

$P_{s,t}$ — the probability of selecting strategy s at time t.

$P_{0,t}$ — the probability of selecting a strategy from S-S(t) at time t.

$w(s,t)$ — the number of times strategy s has been selected up to and including time t.

$\mu_{s,t}$ — an abbreviation for $\mu(A_s(w(s,t)))$, the random variable resulting from the choice of structure made by strategy s at time t.

$P_{S(t)}$ — the probability distribution by which a strategy may be selected from S-S(t) at time t. P is an initial probability distribution over S.

$|S|$ — cardinality of the set S.

$\bar{\mu}_a$ — the mean of the distribution $\mu(a)$ for $a \in A$.

μ_t^* — is the lub $\bar{\mu}_a$ where the lub is over the set $\{a|$ for some strategy $s \in S$, $A_s(t) = a\}$.

$\ddot{\mu}_{s,t}$ — is $\mu_{s,t}$ if strategy s is actually selected at time t and $\mu^*_{w(q,t)}$ if strategy s is not selected but strategy q is selected at time t.

In all of the above notation, if the strategy referred to is in the set $S(t)$, we often use only the subscript. For example we abbreviate $A_{s_j}(k)$ to $A_j(k)$, $w(s_j,t)$ to $w(j,t)$, and $\mu_{s_{j,t}}$ to $\mu_{j,t}$.

We now have enough notation to define a general sequential reproductive plan. There are two calculations which are external to the procedure. The first is the calculation which determines, given the strategy selected by the plan, which structure in A is choosen. In a pure problem basis this calculation always results in the same structure for a given strategy. The second is the calculation of the "payoff" of the structure. These calculations are represented by the functions STRUCTURE and PAYOFF in step 2 of the procedure.

The exact method of selecting new strategies from S-S(t) is not specified in the SRP procedure. We put the following restrictions on the function FIND of step 2 of the procedure. The original probability distribution $P_\emptyset$ over S must be modified to be a distribution $P_{S(t)}$ over $S-S(t)$. We will assume that for $|S|$ finite, if $P_\emptyset(s) > 0$ for a particular strategy s and if $s \notin S(t)$, then $P_{S(t)}(s) \geq P_\emptyset(s)$. For $|S|$ infinite, if $P_\emptyset(W) > 0$ for a particular subset $W \subset S$ and if $W \cap S(t) = \emptyset$, then $P_{S(t)}(W) \geq P_\emptyset(W)$.

Our first procedure is a general form which will be altered in subsequent definitions by changing the variables v and v* and by altering the calculation in step 3.10.

Definition 2.10: SRP, a *sequential reproductive plan*, is a PSAS over a problem basis $\langle A,\mu,S\rangle$ where:

1. The values assumed by the random variables are in an interval (r_1,r_2) where r_1,r_2 are finite real numbers.

2. The procedure requires the following functions:

   2.1 SELECT($\vec{P}_{N(t)}$,j) which assigns a value to j according to the probability vector $\vec{P}_{N(t)}$. That is, if $\vec{P}_{N(t)} = (p_0, p_1, \ldots, p_{N(t)})$ then j is assigned the value i with probability p_i.

   2.2 FIND($P_\emptyset$,S(t),s_j) which chooses a new strategy from S-S(t) according to the restrictions above and labels it s_j.

   2.3 STRUCTURE(s_j,A_j(t),A_0) which applies strategy s_j to obtain a new structure $A_j(t)$.

   2.4 PAYOFF(A_j(t),v,v*) which uses $A_j(t)$ to obtain values for v and v*.

3. The operation of the procedure is as follows:

   3.1 Choose $\Theta > 0$, $k_1 > r_1+1$, $k_2 > 0$, $P_\emptyset$.

   3.2 Choose at random an initial structure A_0.

   3.3 Set N(1) = 0, t = 0, S(t) = $\emptyset$.

   3.4 Set t = t+1.

   3.5 Calculate $\vec{P}_{N(t)} = (P_{0,t}, \ldots, P_{N(t),t})$ by

   3.5.1

$$P_{0,t} = \begin{cases} [N(t)+1]^{-(1+\Theta)} & \text{if } N(t)+1 \le |S| \\ 0 & \text{otherwise} \end{cases}$$

   3.5.2 If N(t) > 0 for $1 \le i \le N(t)$ calculate:

$$P_{i,t} = (1-P_{0,t}) \frac{\mathrm{Prod}_{i,t-1}}{\sum_{h=1}^{N(t)} \mathrm{Prod}_{h,t-1}} .$$

   3.6 SELECT($\vec{P}_{N(t)}$,j)

3.7 If j is 0 then Set $N(t+1) = N(t)+1$, $j = N(t+1)$,

FIND($P_\emptyset$, $S(t)$, s_j), $P_{j,t_j} = P_{0,t}$. $S(t+1) = S(t) \cup \{s_j\}$.

$\mathrm{Prod}_{j,t-1} = P_{j,t_j}$

otherwise set $N(t+1) = N(t)$, $S(t+1) = S(t)$.

3.8 STRUCTURE(s_j, $A_j(t)$, A_0).

3.9 PAYOFF($A_j(t)$, v, v^*).

3.10 Calculate: for $1 \le i \le N(t+1)$

3.10.1
$$\vartheta_i = \begin{cases} v & \text{for } i = j \\ v^* & \text{for } i \neq j \end{cases}$$

3.10.2
$$\mathrm{Prod}_{i,t} = \mathrm{Prod}_{i,t-1} * (\vartheta_i + k_1)^{k_2}.$$

3.11 GOTO step 3.4.

Definition 2.11: An <u>SR1a</u> is an SRP where the value returned by PAYOFF for v is $\mu_{j,t}$ and the value for v^* is $\mu^*_{W(j,t)}$. An <u>SR1b</u> is an SRP where the value returned by PAYOFF for v is $\mu_{j,t}$ and the value for v^* is $\max(\mu^*_{W(j,t)}, \mu_{j,t})$. An SR1 is an SR1a or an SR1b.

The definition of an SR1 differs in several respects from that of Holland (1970). We have allowed the $\mu(a)$ to be random variables. In the original development μ was a single valued function from A to a finite subset of the reals (r_1, r_2). This restriction does not allow the application of the algorithm to such problems as the n-armed bandit problem. In the Holland paper the notion of a set of environments $\mathcal{E}$ was included. However, since all of the theoretical work was done with respect to a fixed element $E \in \mathcal{E}$, this aspect of the plan has been discarded. The notion of environment could easily be incorporated in the functions μ since they are external to the algorithm, or in the structural description of the elements of A. The calculation of the probabilities $P_{i,t}$ in the

SR1 algorithm can be expressed as:

Equation 2.12: $$P_{i,t} = (1-P_{0,t}) \frac{P_{i,t_i} \prod_{t'=t_i}^{t-1} (\bar{\mu}_{i,t'}+k_1)^{k_2}}{\sum_{h=1}^{N(t)} P_{h,t_h} \prod_{t'=t_h}^{t-1} (\bar{\mu}_{h,t'}+k_1)^{k_2}}$$

This is similar to the calculation in the 1970 paper except for the upper limit of the products in the numerator and denominator. In the 1970 paper this limit is t, in equation 2.12 the limit is t-1. We cannot use $\bar{\mu}_{i,t}$ until it has been calculated and in our algorithm this calculation comes in step 3.10 and the calculation of $P_{i,t}$ in step 3.5.

The values of the $P_{0,t}$ determine how often a new strategy is to be tried. One of the weaknesses of this algorithm is that $P_{0,t}$ is not dependent on the performance of the strategies that have already been used. It would be preferable to have $P_{0,t}$ be relatively large if the strategies used so far were not performing well, small if they were doing well. However, this definition does insure that with probability 1, when $|S|$ is finite all of the strategies will be tried by some $T_f < \infty$. When $|S|$ is infinite, there is no finite time after which no new plans are tried.

We observe from equation 2.12 that for an SR1, if $\langle A,\mu,S\rangle$ is deterministic, $P_{j,t}$ cannot increase as a result of using strategy j, except of course, the first time it is used. Also, if the jth strategy is used at time t, $\mu_{j,t} = \mu^*_{w(j,t)}$, and strategy j has been used before, the probabilities $P_{i,t}$ do not change. The purpose of using such a method to change the probability vector over S(t) is to allow the vector tc be more responsive to the performance of the strategy over time. We wish to avoid converging to false strategies due to their early behavior. However, it is exactly this property which makes it impossible for the

algorithm to have all of the convergence properties claimed in Holland (1970).

Let us now consider a modified form of Equation 2.12 for an SRIa procedure for finite S. We know that after some time T_f, each of the strategies will have been used at least once. We also note that the P_{i,t_i} and P_{h,t_h} in the numerator and denominator of Equation 2.12 are constant factors that do not influence the convergence properties of the $P_{j,t}$ as $t \to \infty$. In fact, they reinforce the bias of the original distribution $P_\emptyset$ over S. If this distribution is not a good measure of the value of the strategies, the P_{j,t_j} in the formula could slow down convergence but never prevent it. If we neglect these factors, and let $\bar{\mu}_{j,t} = (\tilde{\mu}_{j,t}+k_1)^{k_2}$ we have:

Equation 2.13: for $t \geq T_f+1$

$$P_{j,t} = \frac{\prod_{n=t_j}^{t-1} (\bar{\mu}_{j,n})}{\sum_{h=1}^{N(t)} \prod_{n=t_h}^{t-1} \bar{\mu}_{h,n}} .$$

Let i be the strategy actually used at time t and $\bar{\mu}^*_t = (\mu^*_{w(i,t)}+k_1)^{k_2}$, Then Equation 2.13 can be expressed as:

Equation 2.14: for $t > T_{f+1}$

if $i = j$:

$$P_{j,t} = P_{j,t-1}\left(\frac{1}{\frac{\bar{\mu}^*_{t-1}}{\bar{\mu}_{j,t-1}} + P_{j,t-1}\left(1 - \frac{\bar{\mu}^*_{t-1}}{\bar{\mu}_{j,t-1}}\right)}\right)$$

if $i \neq j$:

$$P_{j,t} = P_{j,t-1}\left(\frac{1}{1 + P_{i,t-1}\left(\frac{\bar{\mu}_{i,t-1}}{\bar{\mu}^*_{t-1}} - 1\right)}\right)$$

Equation 2.13 has a familiar form. With different interpretations of the $\bar{\mu}$, this is the form of the Beta Model of mathematical psychology. This model has been studied extensively by Luce (1959), Lamperti and Suppes (1960), Lamperti (1960), Norman (1970) and others. Although their terminology and interpretations of the model differ from those of Holland, the mathematical analysis applies to both. The Holland work claims that under certain restrictions of the problem basis $\langle A, \mu, S \rangle$ the model not only converges but converges in such a special way that "bad" strategies are used only a finite number of times. However, the studies by the mathematical psychologists show that the Beta model (without Holland's restrictions) does not always converge.

We will show that under restrictions similar to Holland's the SR1 plan does converge, but not as strongly as he claimed. Then we will examine convergence under other restrictions and relate our results to those of Norman (1970).

CHAPTER 3

DETERMINISTIC PROBLEM BASES

This chapter examines the convergence properties of the SR1 algorithm over deterministic problem bases.

Definition 3.1: A strategy in a deterministic problem basis is *optimal* iff $\exists T \ni \mu(A_s(t)) = \mu_t^*$ for $t > T$.

Definition 3.2: A strategy s in a deterministic problem basis is *asymptotically optimal* iff $\sum_{t=1}^{\infty} [\mu_t^* - \mu(A_s(t))] < \infty$.

If s is optimal, s is asymptotically optimal.

Definition 3.3: An *arena* is a deterministic problem basis <A,μ,S> such that there exists a set $H \subset S$, $P_0(H) > 0$ and $\operatorname*{glb}_{s \in H} \mu(A_s(t)) \geq \mu_t^* - \lambda_t$ where $\sum_{t=1}^{\infty} \lambda_t = N_E < \infty$.

It is not hard to show that <A,μ,S> is an arena if and only if S contains asymptotically optimal strategies with nonzero initial probability.

Definition 3.4: S_E, the *arena set* for the arena <A,μ,S>, is the set of all asymptotically optimal strategies.

To avoid trivial situations we henceforth assume S_E is a proper subset of S. At this point it would appear that if our problem basis is an arena then the optimum algorithm in searching S would converge to the arena set. Because of the lack of limitation on the sequence μ_t^*, this is not necessarily the case. As presently defined, μ_t^* is the best that any single strategy could do if used t times. It is not necessarily the optimum value that any algorithm could achieve at time

t. It is possible that a PSAS which selects an element of S according to a fixed vector $(p_1, p_2, \ldots, p_S)$ with $p_i \neq 0$ for at least one i not in the arena set could have a higher payoff than a similar PSAS with $p_i = 0$ for all strategies not in the arena set. We give an example where this is the case.

Example 3.5: Let strategy 1 have payoff 100/k the kth time it is used, strategy 2 have payoff 99/k the kth time it is used. Then μ_t^* is 100/t and we have an arena set consisting of strategy 1. However, a PSAS that uses strategy 1 all of the time will not have as good a payoff as one that uses an appropriate mix of the strategies 1 and 2. Let p be the probability of choosing strategy 1, 1-p the probability of choosing strategy 2, where p is fixed for all t and $1 > p \geq 1-p$. Then if M is the number of choices of strategy 1 in N trials and $N > M > .01N$, the payoff would be

$$\sum_{k=1}^{M} 100/k + \sum_{k=1}^{N-M} 99/k > \sum_{k=1}^{N} 100/k.$$

For $p > 1-p$, the expected payoff of the mixed strategy is greater than the payoff of the arena set.

To have strategies in a deterministic problem basis have the payoffs indicated, both the set A and the strategies would be quite unnatural. However, such problem bases are not eliminated by the definition of an arena. Consequently we introduce the following notion.

Definition 3.6: An arena $\langle A,\mu,S\rangle$ is a *restricted arena* if there exists a time T such that no PSAS over $\langle A,\mu,S\rangle$ can obtain a payoff at time $t > T$ higher than μ_t^*.

The rest of this section develops the results necessary to show that in fact the losses associated with an SR1 algorithm even over a

restricted arena cannot be finite.

Definition 3.7: The *absolute* loss of strategy s is

$$AL_s = \sum_{t=1}^{\infty} \mu_t^* - \mu(A_s(t)).$$

If a strategy s is in the arena set for <A,μ,S> then AL_s is finite, otherwise AL_s is infinite.

Definition 3.8: The *actual loss* of the strategy s during any particular use of an algorithm is

$$D_{s,t} = (\mu^*_{w(i,t)} - \tilde{\mu}_{s,t})$$

where s_i is the strategy selected at time t. The *accumulated* actual loss is $D_s = \sum_{t=1}^{\infty} D_{s,t}$.

According to Definition 3.8, a strategy s has zero loss at time t if it is not selected at time t. Therefore when we speak of the actual loss incurred by a particular strategy, we only mean the loss resulting from the actual selection of that strategy by an algorithm.

Lemma 3.9: Let S_E be the arena set for an SR1 over a restricted arena <A,μ,S>. If $s \in S_E$, then the accumulated actual loss D_s is finite. If $s' \notin S_E$ and s' is, with probability 1, selected infinitely often, then the accumulated actual loss, $D_{s'}$, is infinite with probability 1.

Proof: The first part of the lemma follows directly fr⌄ the definitions of an arena set and of asymptotically optimal strategies. For the second part observe that:

$$D_{s'} = \sum_{t=1}^{\infty} (\mu^*_{w(i,t)} - \tilde{\mu}_{s',t}) = \sum_{k=1}^{\infty} (\mu^*_{w(s',n_k)} - \tilde{\mu}_{s',n_k})$$

where $\{n_k\}$ is the infinite set of indices where strategy s' is actually

selected. But $w(s',n_k) = k$ so $\tilde{\mu}_{s',n_k} = \mu(A_{s'}(w(s',n_k))) = \mu(A_{s'}(k))$ and

$$D_{s'} = \sum_{k=1}^{\infty}(\mu_k^* - \mu(A_{s'}(k))) = AL_{s'} = \infty \quad \text{for } s' \notin S_E.$$

The definition of actual loss incorporates many of the ideas expressed in Holland (1970). A strategy is to be measured against how well other strategies would have done if they had been used the same number of times. However, there is no guarantee that the actual loss directly represents the loss of using a strategy under SR1 compared with some optimal scheme.

Definition 3.10: The *loss* incurred by an SR1 algorithm over a finite modified arena is

$$D = \sum_{t=1}^{\infty} \sum_{s=1}^{|S|}(\mu_t^* - \tilde{\mu}_{s,t}).$$

Lemma 3.11: The loss of an SR1 algorithm over a finite restricted arena is infinite if the total accumulated actual loss over all strategies is infinite. That is $D = \infty$ if $\sum_{s=1}^{|S|} D_s = \infty$.

Proof: There exists $T \ni$ for $t > T$, $s \in S$, by Definition 3.6

$$\mu_t^* \geqq \mu_{w(i,t)}^* \quad \text{where } s_i \text{ is used at time } t.$$

$$\mu_t^* - \tilde{\mu}_{s,t} \geqq \mu_{w(i,t)}^* - \tilde{\mu}_{s,t}$$

$$\sum_{t=T}^{\infty} \sum_{s\in S} (\mu_t^* - \tilde{\mu}_{s,t}) \geqq \sum_{t=T}^{\infty} \sum_{s\in S}(\mu_{w(i,t)}^* - \tilde{\mu}_{s,t})$$

The left hand side is D-K, the right hand side is $\sum_{s\in S} D_s - K'$ where K and K' are finite constants. Therefore

$$D-K \geqq \sum_{s\in S} D_s - K' = \infty$$

and

$$D = \infty.$$

From Lemma 3.9 the accumulated actual loss is not finite if some strategy s' not in S_E is used infinitely often. We now explore the conditions under which this happens.

Definition 3.12: H is a *limiting set* of a PSAS over the arena $<A,\mu,S>$ if $H \subset S$ and there is some time T such that for $t > T$, with probability 1, only strategies from the set H are selected for use.

If S is finite and there are no proper subsets, $H \subset S$, such that H is a limiting set and $P_\emptyset(H) < 1$, then each of the strategies in S with nonzero initial probability must be used infinitely often.

We want to develop conditions under which a limiting set does not exist. Let Q be a PSAS over $<A,\mu,S>$, $H \subset S$ and $\{X_i\}$ a sequence of random variables defined by

$$X_i = \begin{cases} 1 & \text{if a strategy in H is selected by Q at time i} \\ 0 & \text{otherwise.} \end{cases}$$

We can now represent the choices of strategies of Q as an infinite vector of the X_i. The space of all 0-1 valued infinite vectors is then the underlying space representing all possible sequences of choices by Q. The probability distribution over these vectors is dependent on the vectors $\vec{P}_{N(t)} = \{P_{0,t}, P_{1,t}, \ldots, P_{N(t),t}\}$. In order to have a limiting set, we are interested in those vectors of the X_i which from some point on have entries of 1 only. Let

$$B_m = \{(X_1, X_2, \ldots, X_m, \ldots) \mid X_j = 1 \text{ for all } j \geqq m\}$$

and

$$B = \lim_{m \to \infty} B_m.$$

If the probability of the set B is 0, then H is not a limiting set for

Q. Let

$$\Delta P_m = P(X_{m+1} = 1 | X_1,\ldots,X_{m-1},X_m = 1)$$
$$- P(X_m = 1 | X_1,\ldots,X_{m-1}).$$

Then ΔP_m is the increment in the probability of choosing an element of the set H given that an element in H was chosen at the previous time step, and irrespective of the other past choices. Intuitively, if ΔP_m is negative, one would not expect H to be a limiting set. This is in fact the case.

Theorem 3.13: Let Q be a PSAS over $\langle A,\mu,S\rangle$, $H \subset S$ and X_i, B_m, and B defined as above relative to H and Q. If there is a time T such that, for $m \geq T$, $\Delta P_m \leq 0$, and if $P(X_m = 1) < 1$ for all finite m then $P(B) = 0$ and H is not a limiting set for Q.

Proof: for $m \geq T$,

$$P(B_m) = \lim_{n\to\infty} P(X_i = 1,\ m \leq i \leq n | X_1,\ldots,X_{m-1})$$
$$= \lim_{n\to\infty} P(X_n = 1 | X_1,\ldots,X_{m-1},X_m = 1,\ldots,X_{n-1} = 1)$$
$$\cdot P(X_{n-1} = 1 | X_1,\ldots,X_{m-1},X_m = 1,\ldots,X_{n-2} = 1)$$
$$\cdots P(X_m = 1 | X_1,\ldots,X_{m-1})$$

Since $\Delta P_m \leq 0$,

$$P(B_m) \leq \lim_{n\to\infty} [P(X_m = 1 | X_i,\ldots,X_{m-1})]^{n-m-1}$$
$$= 0$$

So, for arbitrary $m \geq T$, $P(B_m) = 0$.

But $B = \lim_{m\to\infty} B_m$ and the B_m are a sequence of increasing events.

Therefore, $P(B) = \lim_{m\to\infty} P(B_m)$[see I.3.1 Neveu (1965)] so $P(B) = 0$.

Since $P(B)$ is 0, with probability 1 H is not a limiting set of Q.

Lemma 3.14: Let Q be an SR1 algorithm over a restricted arena $\langle A,\mu,S\rangle$ with finite S. Let H be the arena set of S, $H \neq S$ and $P_\emptyset(H) \neq 1$. Then

(a) $P(X_m = 1) < 1$ and (b) $\Delta P_m \geq 0$ for $m > T_f$.

Proof: (a) If $P(X_t = 1) = 1$ then, letting $P_t(H)$ denote the probability of choosing an element of H at t, $P_t(H) = 1$. Therefore $P_t(S\text{-}H) = 0$.

1. If one of the strategies in S-H, say s_j, has been used at least once, $P_t(S\text{-}H) > P_{j,t}$. But $P_{j,t} \neq 0$ if $C_{N(t)+1} \neq 1$ since the $\bar{\mu}_{j,t} > 1$. Since strategy s_j has been used once prior to t, $C_{N(t)+1} \neq 1$ by definition. Therefore $P_t(S\text{-}H) \neq 0$.
2. If none of the strategies in S-H have been used then $P_t(S\text{-}H) = P_{0,t} \cdot P_{S(t)}(S\text{-}H)$. But $P_{S(t)}(S\text{-}H) \geq P_{\emptyset}(S\text{-}H) > 0$. $P_{\emptyset,t} = 0$ only when all strategies have been used. This would imply that $H = S$ and our assumption is that $H \neq S$.

Therefore $P_t(S\text{-}H) \neq 0$ and $P(X_t = 1) < 1$.

(b) Let T_f be the time by which with probability 1 each strategy in S has been used at least once. If strategy $s_i \in H$ is used at time $m > T_f + 1$ then for $j \neq i$ from Equation 2.14 we have,

$$P_{j,m+1} = P_{j,m} \cdot \delta_1 \text{ where } \delta_1 = \frac{1}{1 + P_{i,m}\left(\frac{\bar{\mu}_{i,m}}{\bar{\mu}^*_m} - 1\right)}$$

$$P_{i,m+1} = P_{i,m} \cdot \delta_2 \text{ where } \delta_2 = \frac{1}{\frac{\bar{\mu}^*_m}{\bar{\mu}_{i,m}} + P_{i,m}\left(1 - \frac{\bar{\mu}^*_m}{\bar{\mu}_{i,m}}\right)}$$

since $\delta_1 \geq 1$, $0 < \delta_2 \leq 1$,

$$P_{j,m+1} \leq P_{j,m}, \qquad P_{i,m+1} \geq P_{i,m}$$

Let $\Delta P_{j,m} = P_{j,m+1} - P_{j,m}$, $\quad \Delta P_{i,m} = P_{i,m+1} - P_{i,m}$

Then $\Delta P_{j,m} \geq 0$, $\quad \Delta P_{i,m} \leq 0$

$$\sum_{s_k \in S} \Delta P_{k,m} = \sum_{\substack{s_j \in H \\ j \neq i}} \Delta P_{j,m} + \Delta P_{i,m} + \sum_{s_k \in S-H} \Delta P_{k,m} = 0.$$

$$\sum_{s_j \in H} \Delta P_{j,m} = -\sum_{s_k \in S-H} \Delta P_{k,m} \leqq 0.$$

Therefore, for arbitrary $s_i \in H$ used at time m

$$\Delta P_m = \sum_{s_j \in H} \Delta P_{j,m} \leq 0$$

With probability 1, T_f is finite and therefore, $\Delta P_m \leqq 0$ for all $m > T_f$, T_f finite.

Corollary 3.15: The arena set S_E of a finite arena $\langle A,\mu,S\rangle$ cannot be a limiting set for an SR1. In fact, no $H \subset S$ such that $P_0(H) < 1$ can be a limiting set and each strategy in S with nonzero initial probability $P_\emptyset(s)$ will be used infinitely often by the SR1 algorithm.

We now show that the loss incurred by an SR1 algorithm over a finite restricted deterministic arena is infinite.

Theorem 3.16: With probability 1, the loss incurred by an SR1 algorithm over a finite restricted arena $\langle A,\mu,S\rangle$ is not finite. That is with probability 1,

$$\sum_{s=1}^{|S|} D_s = \infty$$

and therefore

$$D = \infty.$$

Proof: By Corollary 3.15, the arena set S_E of $\langle A,\mu,S\rangle$ is not a limiting set. Therefore with probability 1, at least one strategy $s' \notin S_E$ is used infinitely often. Therefore Lemma 3.9 shows that for s', with

probability 1,

$$\sum_{t=1}^{\infty} D_{s',t} = \infty$$

since $\quad D_{s,t} = (\mu^*_{w(i,t)} - \tilde{\mu}_{s,t}) \geq 0$ for all s,t,i

$$\sum_{s=1}^{|S|} D_s = \sum_{t=1}^{\infty} \sum_{s=1}^{|S|} D_{s,t} > \sum_{t=1}^{\infty} D_{s',t} = \infty$$

By Lemma 3.11, $\sum_{s=1}^{|S|} D_s = \infty \Rightarrow D = \infty$.

We have shown now that an SR1 over a restricted deterministic arena will, with probability 1, not achieve finite losses. However, the theorem in Holland (1970) is interpreted as showing that a reproductive plan would, with probability 1, achieve finite losses. Let us define the "expected" loss of the algorithm to be:

3.17 $$EL_t = \mu^*_t - \sum_{k=1}^{N(t)} \tilde{\mu}_{k,t} P_{k,t} - P_{0,t} \overline{X}_{S-S(t)}$$

where $\overline{X}_{S-S(t)}$ is the expected payoff resulting from strategies in S-S(t) and the $P_{k,t}$ are defined as in 2.10. Now Holland's result in our notation states:

Proposition 3.18: (Holland) An SR1 algorithm over a restricted arena $\langle A,\mu,S\rangle$ will, with probability 1, satisfy the following criteria:

(1) $\lim_{T\to\infty} \sum_{t=1}^{T} EL_t < \infty$ if $|S|$ is finite;

(2) $\lim_{T\to\infty} (\sum_{t=1}^{T} \mu^*_t - \sum_{t=1}^{T} EL_t)/ \sum_{t=1}^{T} \mu^*_t = 1$ otherwise.

This proposition would appear to conflict with our results. However, this expression for expected loss, EL_t is less than the loss the algorithm would actually receive at any time t. That is for $|S|$

finite and $t > T_f$

$$EL_t = \mu_t^* - \sum_{k=1}^{|S|} \mu_{k,t} P_{k,t} = (\mu_t^* - \mu_{j,t}) P_{j,t} \qquad \text{strategy } s_j \text{ used at time } t.$$

$$< \mu_t^* - \mu_{j,t}.$$

Therefore, showing that $\sum_{t=1}^{\infty} EL_t$ is finite in no way implies that any actual use of the algorithm will achieve finite losses, and EL_t is not an adequate expression for the expected loss of the algorithm.

We could change the definition of an SR1 algorithm so that the strategies in S were used in a more parallel fashion. That is, the strategy chosen at time t actually would calculate the element of A it would calculate if it were being used for the t th time. If we define

$$y_{k,t} = \begin{cases} \mu(A_k(t)) & \text{if } s_k \text{ is used at time } t \\ \mu_t^* & \text{otherwise} \end{cases}$$

$y_{k,t}$ represents the evaluation a strategy would receive if it had actually been used t times, whether it had or not. This would be possible if the strategies in S did not use the "feedback" evaluation, $\mu(A_s(t-1))$, of the element of A selected at time t-1 to decide which element of A to select at time t.

Definition 3.19: An SR2 is an SRP where the function PAYOFF returns the value $\mu(A_j(t))$ for v and μ_t^*.

The expected loss defined analogous to 3.17,

3.20 $$EL_t' = \mu_t^* - \sum_{k=1}^{N(t)} y_{k,t} P_{k,t} - P_{0,t} \overline{X}_{S-S(t)}$$

is still less than the actual loss on any particular use of the algorithm. However, we cannot guarantee that the loss would be

infinite. Even though each plan will be used infinitely often, it is possible that for some $s' \notin S_E$, $\{k_n\}$ the infinite set of times at which s' is used,

$$\sum_{n=1}^{\infty} (\mu^*_{k_n} - y_{s',k_n}) = \sum_{n=1}^{\infty} (\mu^*_{k_n} - \mu(A_{s'}(k_n))) < \infty .$$

If this happens for all strategies in S, then the loss will be finite. We see immediately, however, that the loss is not finite if even one strategy in S has an evaluation strictly bounded away from μ^*_t.

We finish this chapter by demonstrating that under the restrictions we have been imposing, an SR1 algorithm will converge, with probability 1, to a subset of the arena set of S.

Theorem 3.21: An SR1 over a finite, restricted deterministic arena $\langle A,\mu,S\rangle$ converges with probability 1 to a set $H \subseteq S_E$, where S_E is the arena set of S. That is, with probability 1,

$$\lim_{t\to\infty} \sum_{s_j \in H} P_{j,t} = 1.$$

Proof: By Corollary 3.15, every strategy in S with a nonzero initial probability must be used infinitely often by the SR1 algorithm. Consider the products

$$\Gamma_{j,t-1} = \prod_{n=t_j}^{t-1} \frac{\bar{\mu}_{j,n}}{\bar{\mu}^*_n} = \prod_{n=t_j}^{t-1} \frac{\bar{\mu}_{j,n}+k_1}{\mu^*_{w(\ell,n)}+k_1}$$

if $\ell \neq j$, the numerator and the denominator are equal. If m_q, $q=1,\ldots,M$, are the times at which $\ell = j$; that is, the times at

which strategy j is actually used, then $w(j,m_q) = q$ and

$$\epsilon_{j,t-1} = \prod_{q=1}^{M} \frac{\mu(A_j(q))+k_1}{\mu_q^*+k_1}$$

Then
$$\lim_{t\to\infty} \epsilon_{j,t-1} = \lim_{M\to\infty} \prod_{q=1}^{M} \frac{\mu(A_j(q))+k_1}{\mu_q^*+k_1}.$$

Now,
$$\frac{\mu(A_j(q))+k_1}{\mu_q^*+k_1} = 1 - \frac{\mu_q^*-\mu(A_j(q))}{\mu_q^*+k_1}$$

for $s_j \in S_E$
$$\sum_{q=1}^{\infty} \frac{\mu_q^*-\mu(A_j(q))}{\mu_q^*+k_1} < \sum_{q=1}^{\infty} [\mu_q^*-\mu(A_j(q))] < \infty$$

for $s_j \notin S_E$
$$\sum_{q=1}^{\infty} \frac{\mu_q^*-\mu(A_j(q))}{\mu_q^*+k_1} \geq \frac{1}{v_2+k_1} \sum_{q=1}^{\infty} [\mu_q^*-\mu(A_j(q))] = \infty.$$

We now recall the following basic result [Theorem 7, page 96 of Knopp (1956)]:

A product of the form $\prod_{v=1}^{\infty}(1-a_v)$, with $0 \leq a_v \leq 1$, is convergent (to a nonzero value) if, and only if, $\sum_{v=1}^{\infty} a_v$ converges.

Applying this we find that

(a) for $s_j \in S_E$ $\lim_{t\to\infty} \epsilon_{j,t-1} = \lim_{M\to\infty} \prod_{q=1}^{M} \frac{\mu(A_j(q))+k_1}{\mu_q^*+k_1} = N_{\epsilon,j} > 0$

(b) for $s_j \notin S_E$ $\lim_{t\to\infty} \epsilon_{j,t-1} = \lim_{M\to\infty} \prod_{q=1}^{M} \frac{\mu(A_j(q))+k_1}{\mu_q^*+k_1} = 0$

Now
$$\lim_{t\to\infty} P_{j,t} = \lim_{t\to\infty} \frac{P_{0,\tau_j} \prod_{n=1}^{t_j} \frac{1}{\mu_n^*} \cdot \epsilon_{j,t-1}}{\sum_{\ell=1}^{|S|} P_{0,t_\ell} \prod_{n=1}^{t_\ell} \frac{1}{\mu_n^*} \cdot \epsilon_{\ell,t-1}}$$

The denominator is always nonzero since we are in an arena and S_E is not empty.

from (b) if $s_j \notin S_E$ $\lim_{t \to \infty} P_{j,t} = 0$

from (a) if $s_j \in S_E$ $\lim_{t \to \infty} P_{j,t} = \dfrac{P_{0,t_j} N_{\varepsilon,j} \prod_{n=1}^{t_j} \frac{1}{\mu_n^*}}{\sum_{\ell \in S_E} P_{0,t_\ell} N_{\varepsilon,\ell} \prod_{n=1}^{t_\ell} \frac{1}{\mu_n^*}}$

With probability 1, the times t_ℓ are finite for those $\ell \in H \subset S_E$ such that $P(\ell) > 0$ and therefore $\lim_{t \to \infty} P_{j,t} \neq 0$.

Therefore, $\lim_{t \to \infty} \sum_{j \in S_E} P_{j,t} = 1$.

Now we have shown that an SR1 algorithm converges to a set of good strategies in a very restricted problem basis. However, the proof of Theorem 3.20 does not generalize to a problem basis that is not an arena. This is obvious because if <A,μ,S> is not an arena, we cannot guarantee the convergence for any of the $\varepsilon_{j,t}$. The proof also assumes that $\mu_t^* > \mu(A_j(t))$. This assumption is not always justified in a nondeterministic problem basis.

CHAPTER 4

PURE PROBLEM BASES

Section I: The Search for an Arena

Let us now consider a pure problem basis, $\langle A,\mu,S\rangle$, where S is a set of pure strategies and the variances of the random variables $\mu(a)$, $a \varepsilon A$ are non-zero. We suppose there is a one-to-one, onto mapping between A and S. We will use ρ_s for ρ_a, the mean of the function $\mu(a)$. We will denote $\mu(a)$ by μ_s. We note that $\mu^* = \underset{s\varepsilon S}{\text{lub}}\ \rho_s$ is not dependent on the time variable since we are dealing with pure strategies in a nonvariable problem basis.

Let us propose some definitions for asymptotically optimal strategies in this basis. If we can find such definitions we can then define an arena and an arena set over a pure problem basis similar to Definitions 3.3 and 3.4 for a deterministic problem basis. If not, we must find other ways to explore the convergence of SR type algorithms over pure problem bases. The two most obvious candidates for such a definition are:

A. A strategy $s \varepsilon S$ in a pure problem basis $\langle A,\mu,S\rangle$ is asymptotically optimal if $\left|\lim_{n\to\infty} \sum_{t=1}^{n} (\mu^*-\mu_{s,t})\right| < +\infty$ a.s.

B. A strategy $s \varepsilon S$ in a pure problem basis $\langle A,\mu,S\rangle$ is asymptotically optimal if $\left|\lim_{n\to\infty} \sum_{t=1}^{n} (\nu_t^*-\mu_{s,t})\right| < +\infty$ a.s. where $\nu_t^* = \max_{s\varepsilon S} (\mu_{s,t},\mu^*)$.

If we accept either of these definitions, there are no asymptotically optimal strategies in a pure problem basis. We prove this statement in the following theorems.

Lemma 4.1: Let $\{Y_t\}$ be a sequence of independent identically distributed real random variables, then

a. If $\exists\ \epsilon > 0$, $\delta > 0$, N such that for all $t \geq N$, $P(Y_t > \epsilon) > \delta$, then $\lim_{n\to\infty} \sum_{t=1}^{n} Y_t$ a.s. does not exist.

b. If $\exists\ \epsilon < 0$, $\delta > 0$, N such that for all $t \geq N$, $P(Y_t < \epsilon) > \delta$, then $\lim_{n\to\infty} \sum_{t=1}^{n} Y_t$ a.s. does not exist.

Proof: (a) from the hypothesis $\sum_{t=1}^{\infty} P(Y_t > \epsilon) = +\infty$. Since the Y_t are independent identically distributed real random variables (iidrrv), the Borel-Cantelli theorem tells us that with probability 1, $Y_t > \epsilon$ infinitely often. Therefore the terms of the sequence $\{Y_t\}$ do not go to zero and the $\lim_{n\to\infty} \sum_{t=1}^{n} Y_t$ a.s. does not exist. The proof of part b is similar.

Theorem 4.2: There are no strategies in a pure problem basis which are asymptotically optimal according to definition A.

Proof: Let $Y_{s,t} = (\mu^* - \mu_{s,t})$. Then $E[Y_t] = \mu^* - \rho_s \geq 0$, $VAR[Y_t] = VAR[\mu_s] > 0$ and the $Y_{s,t}$ are iidrrv. Therefore, there is an ϵ, δ and an N satisfying part 1 of Lemma 4.1 and the theorem follows immediately.

Theorem 4.3: There are no strategies in a pure problem basis which are asymptotically optimal according to definition B.

Proof: Let $Y_{s,t} = (\nu_t^* - \mu_{s,t})$. By definition of ν_t^*, $Y_{s,t} \geq 0$. Since the $Y_{s,t}$ are iidrrv and $VAR[\mu_s] = \delta > 0$, $VAR[Y_{s,t}] = \delta_1' > 0$ and we can again apply part (a) of the lemma to show that $\lim_{n\to\infty} \sum_{t=1}^{n} Y_{s,t}$ a.s. does not exist. In fact, since all of the terms in the sequence are non-negative, $\lim_{n\to\infty} \sum_{t=1}^{n} Y_{s,t} = +\infty$ a.s.

These results are intuitive and essentially just illustrate that the $\mu_{s,t}$ are independent samples of the real random variable μ_s.

Theorems 4.2 and 4.3 show that the concept of asymptotically optimal strategies in a pure problem basis as defined by A and B is vacuous.

We now consider using the sample mean as the payoff of a strategy as suggested in Holland (1970). This may be done in two ways. We can require that the algorithm perform the calculation of the sample mean. The algorithm then has the additional task of storing the necessary information to calculate the sample mean for each strategy. Alternatively we could change to a variable problem basis where the calculation of the sample mean is made outside the algorithm in the evaluation of the function payoff. The algorithm would treat the sample mean as the "payoff" for the strategy. The exact implementation is not critical to the question under investigation: Are there asymptotically optimal strategies under Definitions A or B when $\mu_{s,t}$ is replaced by the sample mean for strategy s? We will formulate a new version of the SR algorithm which calculates the sample mean and uses this value to calculate the probability vector over the strategy set.

Definition 4.4: An SR3A (SR3B) is an SRP with the following changes: PAYOFF returns the value $\mu_{j,t}$ for v, μ^* for v^*. In step 3.7 of the definition of an SRP (Def. 2,10),if j is 0 set $\lambda_{j,0} = 0$. Step 3.10 of definition 2.10 is replaced by:

3.10 Calculate:

1. $$\lambda_{j,w(j,t)} = \frac{\lambda_{j,w(j,t)-1}.(w(j,t)-1)+\mu_{j,t}}{w(j,t)}$$

2. for $1 \leq i \leq N(t+1)$

$$\hat{v}_{i,t} = \begin{cases} \lambda_{j,w(j,t)} & i = j \\ v^* & i \neq j \end{cases}$$

We define SR3B as above except step 3.10.2 is:

$$\nu_{i,t} = \begin{cases} \lambda_{j,w(j,t)} & i = j \\ \max(\nu^*, \lambda_{j,w(j,t)}) & i \neq j \end{cases}$$

The algorithms SR3A and SR3B suggest the following definitions for asymptotically optimal strategies:

C. A strategy $s \in S$, in a pure problem basis $\langle A,\mu,S\rangle$ is asymptotically optimal if $\left|\lim_{n\to\infty} \sum_{t=1}^{n} (\mu^*-\lambda_{s,t})\right| < +\infty$ a.s.

D. A strategy $s \in S$ in a pure problem basis $\langle A,\mu,S\rangle$ is asymptotically optimal if $\left|\lim_{n\to\infty} \sum_{t=1}^{n} (\nu_t^*-\lambda_{s,t})\right| < +\infty$ a.s. where $\nu_t^* = \max(\mu^*,\lambda_{s,t})$.

We now show that definition C cannot be satisfied by any strategy in a pure problem basis.

Lemma 4.5: Let $\langle A,\mu,S\rangle$ be a pure problem basis. For those strategies $s \in S$ such that $\rho_s \neq \mu^*$, $\lim_{n\to\infty} \sum_{t=1}^{n} (\mu^*-\lambda_{s,t})$ a.s. does not exist.

Proof: $E[\mu^*-\lambda_{s,t}] = \mu^*-\rho_s > 0$. Choose ε so that $\mu^*-\rho_s > \varepsilon > 0$, then $P((\mu^*-\lambda_{s,t}) > \varepsilon) \geq \delta$ for $\delta > 0$ and Lemma 4.1 a. shows that $\lim_{n\to\infty} \sum_{t=1}^{n} (\mu^*-\lambda_{s,t})$ a.s. does not exist.

Lemma 4.5 shows that strategies which do not obtain the maximum mean cannot be asymptotically optimal. The next lemma shows that even if strategies do obtain the maximal mean they cannot by asymptotically optimal by definition C.

Lemma 4.6: Let $\mu_1,\mu_2,\ldots,\mu_t,\ldots$ be independent identically distributed real random variables with nonzero variance.

Let $E[\mu_t] = \mu^*$, $\lambda_t = \frac{\sum_{i=1}^{t} \mu_i}{t}$. Then it is not the case that

$\left|\lim_{n \to \infty} \sum_{t=1}^{n} (\mu^* - \lambda_t)\right| < \infty$ with probability 1.

Proof: Let Y_n be the n^{th} partial sum. Then rearranging terms

$$Y_n = (\mu^* - \mu_1) \sum_{i=1}^{n} 1/i + \ldots + (\mu^* - \mu_n)\ 1/n.$$

$$\text{Let } X_n = (\mu^* - \mu_1) \sum_{i=1}^{n} 1/i\ ;\quad Z_n = Y_n - X_n.$$

X_n and Z_n are independent random variables since the μ_t are independent.

Since $\sum_{i=1}^{\infty} 1/i$ does not converge and the variance of μ_t is nonzero, for any M there is an N_1, and δ_1, such that for $n > N_1$

$$P(X_n > 2M) > \delta_1$$

$$P(X_n < -2M) > \delta_1 \qquad \text{where } 2\delta_1 < 1.$$

Suppose that the $\lim_{m \to \infty} Y_m = N < \infty$ with probability 1. Then for $\varepsilon = 1/3\delta_1$, there is an N_2 such that for $n > N_2$

$$P(|Y_n| > M) < \varepsilon = (1/3)\delta_1$$

Let $N = \max(N_1, N_2)$. There are two mutually exclusive events (not exhaustive) that give $|Y_m| > M$ namely, $\{X_m > 2M \text{ and } Z_m > -M\}$ and $\{X_m < -2M \text{ and } Z_m < M\}$.

Therefore we know that

$$P(X_m > 2M,\ Z_m > -M) < (1/3)\delta_1$$

$$P(Z_m > -M) < (1/3)\delta_1 / P(X_m > 2M) < 1/3$$

$$P(Z_m \le -M) > 2/3$$

Similarly:

$$P(Z_m \ge +M) > 2/3$$

Since $\{Z_m \le -M\}$ and $\{Z_m \ge +M\}$ are exclusive events, this is impossible. Therefore the assumption about the convergence of $\lim_{n\to\infty} \sum_{t=1}^{n} (\mu^* - \lambda_t)$ is false and the theorem results.

Theorem 4.7: There are no strategies in a pure problem basis which are asymptotically optimal according to definition C.

Proof: The proof follows immediately from Lemma 4.5 and 4.6.

We would like an even stronger result than Theorem 4.7. We would like to be able to say that the limit in definition C is nonconvergent with probability 1 even for strategies with optimal mean. Similarly, we would like to be able to say that the limit in definition D is nonconvergent with probability 1. The following two theorems are the result of collaboration with B. Koopmann, C. Qualls, P. Pathok. The proofs are contained in Appendix A.

Theorem 4.8: Let $\{\mu_t\}$ be a sequence of independent identically distributed random variables with nonzero variance and finite second moments, let $E[\mu_t] = \mu^*$, then $\lim_{n\to\infty} \sum_{t=1}^{n} (\nu_t^* - \lambda_t) \to \infty$ a.s. where

$$\lambda_t = \sum_{i=1}^{t} \frac{\mu_i}{t}, \quad \nu_t^* = \max(\mu^*, \lambda_t).$$

Theorem 4.9: Let $\{\mu_t\}$ be a sequence of independent identically distributed random variables with nonzero variance and let $E[\mu_t] = \mu^*$, then one of the following three conditions hold:

i) $\sum \frac{S_n}{n} \to +\infty$ a.s. ii) $\sum \frac{S_n}{n} \to -\infty$ a.s. or

iii) $\underline{\lim} \sum \frac{S_n}{n} \to -\infty$ a.s., $\overline{\lim} \sum \frac{S_n}{n} \to -\infty$ a.s. In all three cases $\sum \frac{S_n}{n}$ diverges a.s. where $S_n = \sum_{t=1}^{n} \mu_t$.

We can now obtain the theorem we need concerning definition D:

Theorem 4.10: There are no strategies in a pure problem basis which are asymptotically optimal according to definition D.

Proof: Theorem 4.8 shows that strategies having optimal means ($\rho_s = \mu^*$) cannot be asymptotically optimal according to definition D. Let s be a strategy with $\rho_s < \mu^*$, then

$$\nu_t^* - \lambda_{s,t} \geq \hat{\rho}_t - \lambda_{s,t} \qquad \text{for all } t,$$

where $\nu_t^* = \max(\mu^*, \lambda_{s,t})$, $\hat{\rho}_t = \max(\rho_s, \lambda_{s,t})$.

$$\text{Therefore} \quad \lim_{n\to\infty} \sum_{t=1}^{n} (\nu_t^* - \lambda_{s,t}) \geq \lim_{n\to\infty} \sum_{t=1}^{n} (\hat{\rho}_t - \lambda_{s,t}) \to +\infty \quad \text{a.s.}$$

Thus no strategies can be asymptotically optimal according to D.

We have shown that using knowledge of the true means of the distribution of the μ_j does not provide a definition for asymptotic optimality. However, we could use an estimate of the means as a value for μ^*. Let f be a function of the sample means, we want to know if f can be defined so that for strategies t_j with $\rho_j = \mu^*$

$$\sum_{t=1}^{\infty} [f(\lambda_{1,t}, \ldots, \lambda_{|S|,t}) - \lambda_{j,t}] < \infty \quad \text{a.s.}$$

and for strategies t_i with $\rho_i < \mu^*$

$$\sum_{t=1}^{\infty} [f(\lambda_{1,t}, \ldots, \lambda_{|S|,t}) - \lambda_{i,t}] = \infty \quad \text{a.s.}$$

We have been exploring this problem but have as yet not developed a function satisfying these conditions.

In order to examine the convergence of SRP algorithms over pure problem bases, we conclude that we must employ other means than we used for deterministic bases. The next section introduces the notion of Linear Additive Models, and in Section III of this chapter we show that in most cases the SRP procedures do not converge in a pure problem bases.

Section II: The Linear Additive Model

In this section we will summarize the development of the Linear Additive Model, a generalization of the Beta Model, given by Norman [1970]. It should be noted that Norman's results are only for the case $|S| = 2$. This may appear to be a very limited analysis. However, all of the major problems of convergence can be found in this "two dimensional" case. In the psychological interpretation of the model, one is usually interested in the two choice situation. Lamperti [1960] has obtained some results for the Beta Model for the general case. Several of the results presented in this section will be used later on. In Section III we relate Norman's assumptions to the SRP plans we have defined.

Assumptions (Norman):

4.11 Q_i, $i = 1,2$, is a probability distribution over the Borel subsets of R, such that $M_i(\zeta) = \int_{-\infty}^{\infty} e^{\zeta x} dQ_i(x)$ exists for ζ in some open interval J containing 0.

4.12 p is a measurable mapping of R into $I = [0,1]$, such that $p(L) \to 1$ as $L \to \infty$, $p(L) \to 0$ as $L \to -\infty$, and p is bounded away from 0 and 1 on any finite interval. Let $q = 1-p$.

4.13 $T_t = (L_t, X_t)$, $t \geq 1$, is a bivariate stochastic process in $R \times \{1,2\}$, such that

$$P(X_t = 1 | L_t, T_{t-1}, \ldots, T_1) = p_t \quad (p_t = p(L_t)),$$
$$P(X_t = 2 | L_t, T_{t-1}, \ldots, T_1) = q_t \quad (q_t = q(L_t)).$$

and for any B

$$P(\Delta L_t \in B | T_t, T_{t-1}, \ldots, T_1) = Q_{X_t}(B) \text{ almost surely.}$$

The means of the conditional distributions of L_t are very important. In fact, their values determine the convergence or divergence of the sequence L_t. Let $m_i = \int_{-\infty}^{\infty} x dQ_i(x)$, then $m_i = M_i'(0)$, for $i = 1,2$.

In order for the sequence p_t to converge, L_t must be absorbed at either $+\infty$ or $-\infty$. Lemmas 4.14 and 4.15 state conditions for absorption and reflection if the m_i are nonzero and are constant. Lemma 4.16 removes the possibility for convergence of L_t to some finite limit, with nonzero m_i.

Lemma 4.14: (Norman) a. If $m_1 < 0$, $\lim\inf L_t < \infty$ a.s. For any $\delta > 0$ such that $M_1(\delta) < 1$, there is a constant $B(\delta)$ such that $\lim\sup E_L[v_t^\delta] \leq B(\delta)$.

b. If $m_2 > 0$, $\lim\sup L_n > -\infty$ a.s. For any $\delta < 0$, such that $M_2(\delta) < 1$, there is a constant $B(\delta)$ such that $\lim\sup E_L[v_t^\delta] \leq B(\delta)$.

Lemma 4.15: (Norman) a. If $m_1 > 0$, then $\lim\sup L_t = \infty$ implies $\lim L_t = \infty$ a.s., and $g_0(L) = P_L(\lim L_n = -\infty) \to 0$ as $L \to \infty$.

b. If $m_2 < 0$, then $\lim\inf L_t = -\infty$ implies $\lim L_t = -\infty$ a.s. and $g_1(L) = P_L(\lim L_t = \infty) \to 0$ as $L \to -\infty$.

Lemma 4.16: (Norman) a. If $m_1 > 0$ or $m_2 > 0$, $P(\lim\sup L_t \in R) = 0$.

b. If $m_1 < 0$ or $m_2 < 0$, $P(\lim\inf L_t \in R) = 0$.

With these lemmas, four possibilities for the asymptotic behavior of p_t can be distinguished provided that the m_i are nonzero and that the Q_i are not dependent upon t. Either p_t converges to 1 a.s., converges to 0 a.s., does not converge to any limit, or has a probability π of converging to 0 and a probability $1-\pi$ of converging to 1.

Theorem 4.17. (Norman)

a. If $m_1 > 0$ and $m_2 > 0$, $\lim p_t = 1$ a.s.

b. If $m_1 < 0$ and $m_2 < 0$, $\lim p_t = 0$ a.s.

c. If $m_1 > 0$ and $m_2 < 0$, $g_1(L) + g_2(L) = 1$, where $g_1(L) = P_L(\lim p_t = 1)$,

$g_2(L) = P_L(\lim p_t = 0)$. In addition $g_1(L) > 0$,

$g_2(L) > 0$, $g_1(L) \to 1$ as $L \to +\infty$ and $g_2(L) \to 1$ as $L \to -\infty$.

d. If $m_1 < 0$ and $m_2 > 0$, $\lim\sup p_t = 1$ and $\lim\inf p_t = 0$ a.s.

These results are independent of the initial value of L_1.

Variations of this theorem for the Beta Model can be found in Lamperti and Suppes [1960].

Two relations are developed by Norman which will prove useful later: Let $v_t = e^{Lt}$, then

Equation 4.18:

$$E[(v_{t+1}|v_t)^\lambda|v_t] = M_1(\lambda)p_t + M_2(\lambda)q_t$$

Equation 4.19:

$$E[v_{n+1}^\lambda|v_n] = v_n^\lambda[M_1(\lambda)p_n + M_0(\lambda)q_n]$$

Section III: Linear Models and Pure Problem Bases

Now we relate Norman's assumptions to the SR1 algorithm where $|S| = 2$. Let T_f be the time at which the second strategy is selected. This time is finite with probability one.

Let

$$L_{T_f} = \sum_{t=1}^{T_f-1} \ln(\mu_{1,t}+k_1)$$

$$L_t = L_{t-1} + \ln\frac{\mu_{1,t-1}+k_1}{\mu_{2,t-1}+k_1} \qquad \text{for } t > T_f \quad .$$

Then from 2.12 and 2.13 we obtain:

Equation 4.20

$$\text{for } t \le T_f \qquad P_{1,t} = 1-2^{-(1+\theta)}$$

$$P_{2,t} = 2^{-(1+\theta)}$$

$$\text{for } t > T_f \qquad P_{1,t} = \frac{2^{(1+\theta)}e^{k_2 L_t}}{2^{(1+\theta)}e^{k_2 L_t}+1} = p(L_t)$$

$$P_{2,t} = 1-p(L_t) = q(L_t) \quad .$$

For $t > T_f$, assumption 4.12 of Norman is certainly satisfied by equation 4.20. Let $\Delta L_t = L_{t+1} - L_t$ and $Q_{i,t}$ be the probability distribution of $\mu_{i,t}$. If strategy i is used at time t we use the notation $\Delta L_{i,t}$. Let

Equation 4.21. $\quad M_{i,t}(\zeta) = \int_{-\infty}^{\infty} e^{\zeta\Delta L_{i,t}}\, dQ_{i,t}(\mu_{i,t}),\ i= 1,2.$

Since $\mu_{i,t} + k_1 \ge 1$ and by the assumption that the moment generating functions of the $\mu(a)$ exist (definition 2.1) $M_{i,t}(\zeta)$ exists for ζ in some open interval J containing 0. Therefore, assumption 4.11 is satisfied if we replace $Q_{i,t}$ with Q_i. If the set S of strategies is a set of pure strategies then we can remove the dependence on the time variable and satisfy assumption 4.11. Assumption 4.13 follows from the

definitions of $P_{i,t}$, L_t and Q_i in a straight forward manner.

Unfortunately, SRP plans do not fall simply into the categories for convergence outlined by Norman. Therefore, we must first extend his results to include the cases where the m_i might be 0.

Lemma 4.22. If $m_1 = 0$, $\sigma_1 > 0$, or $m_2 = 0$, $\sigma_2 > 0$, then
(a) $P(\limsup L_t \in R) = 0$. (b) $P(\liminf L_t \in R) = 0$.
Proof: Norman's proof of Lemma 4.16 relies on the observation that if $m_i > 0$, then $Q_i([2\varepsilon,\infty)) > 0$ for some $\varepsilon > 0$ and $Q_i([2\varepsilon,-\infty)) > 0$ for some $\varepsilon < 0$. This same observation holds if $m_i = 0$ and we know that $\sigma_i > 0$. Once we have made this observation, Norman's proof applies to the present lemma.

Lemma 4.23. a) if $m_1 = 0$, $\sigma_1 \neq 0$ then $\lim L_t \neq +\infty$.
b) if $m_2 = 0$, $\sigma_2 \neq 0$ then $\lim L_t \neq -\infty$.

Proof:

a) Suppose $\lim L_t = +\infty$ a.s. Since $m_1 = 0$, we can chose a $\lambda < 0$ and an $\varepsilon > 0$ such that $M_1(\lambda) > 1+\varepsilon$. Then

$$\lim_{t\to\infty} \lambda L_t = -\infty \quad ,$$

$$\lim_{t\to\infty} e^{\lambda L_t} = 0 \quad ,$$

and

$$\lim_{t\to\infty} [M_1(\lambda)\, p_t + M_2(\lambda)\, q_t] = M_1(\lambda) \quad .$$

Given ε we can find N_o such that for all $t > N_o$

$$M_1(\lambda)\, p_t + M_2(\lambda)\, q_t > M_1(\lambda) - \varepsilon$$

by Equation 4.19

$$E[e^{\lambda L_{t+1}}] \geq (M_1(\lambda)-\varepsilon)^{t-N_o-1}(\min[M_1(\lambda) \quad ,M_2(\lambda)])^{N_o}\cdot e^{\lambda L_1}$$

$$\lim_{t\to\infty} E[e^{\lambda L_{t+1}}] > K \lim_{t\to\infty} (M_1(\lambda) - \varepsilon)^{t-N_o-1} \quad \text{where } k > 0.$$

But $\lim E[e^{\lambda L_{t+1}}] = 0$ so $K \cdot \lim (M_1(\lambda) - \varepsilon)^{t-N_0-1} \leq 0$

By our choice of λ and ε, $M_1(\lambda) - \varepsilon$ is a constant greater than 1.

Since K is positive, the limit cannot be less than or equal to 0.

Therefore we have a contradiction and our original assumption that

$L_t = \infty$ is incorrect. The proof of b) is similar.

Theorem 4.24

a) If $m_1 = 0$, $m_2 > 0$, $\sigma_i > 0$ then $\lim\sup p_t = 1$ and $\lim\inf p_t = 0$.

b) If $m_1 < 0$, $m_2 = 0$, $\sigma_i > 0$ then $\lim\sup p_t = 1$ and $\lim\inf p_t = 0$.

c) If $m_1 = 0$, $m_2 = 0$, $\sigma_i > 0$ then $\lim\sup p_t = 1$ and $\lim\inf p_t = 0$.

d) If $m_1 = 0$, $m_2 < 0$, $\sigma_i > 0$ then $\lim p_t = 0$ a.s.

e) If $m_1 > 0$, $m_2 = 0$, $\sigma_i > 0$ then $\lim p_t = 1$ a.s.

Proof:

a) From lemma 4.14b, $\lim\sup L_t > -\infty$a.s., lemma 4.22a shows that $p(\lim\sup L_t \varepsilon R) = 0$. Therefore $\lim\sup L_t = +\infty$a.s. and by 4.12, $\lim\sup p_t = 1$.

From Lemma 4.22b $P(\lim\inf L_t \varepsilon R) = 0$. If $\lim\inf L_t - +\infty$a.s., then since $\lim\sup L_t = +\infty$a.s., $\lim L_t = +\infty$a.s. But since $\sigma_1 \neq 0$ this contradicts Lemma 2.23a. Therefore $\lim\inf L_t = -\infty$a.s, and by 4.12 $\lim\inf p_t = 0$.

b) proof is similar to a.

c) By lemma 4.22, $P(\lim\sup L_t \varepsilon R) = 0$, $P(\lim\inf L_t \varepsilon R) = 0$. By lemma 4.23 $\lim L_t \neq +\infty$ and $\lim L_t \neq -\infty$. Therefore, $\lim\sup L_t \neq \lim\inf L_t$ and $\lim\sup L_t = +\infty$a.s., $\lim\inf L_t = -\infty$a.s. and the theorem follows.

d) By lemma 4.22, $\lim\inf L_t \notin R$ a.s. and $\lim\sup L_t \notin R$ a.s. If $\lim\inf L_t = +\infty$ then $\lim\sup L_t = +\infty$ and $\lim L_t = +\infty$. But this contradicts lemma 4.23a. Therefore $\lim\inf L_t = -\infty$ and by lemma 4.15b, and $m_2 < 0$, $\lim L_t = -\infty$a.s., and by assumption 4.12 $\lim p_t = 0$ a.s.

e) By lemma 4.22, lim inf $L_t \neq R$ a.s. and lim sup $L_t \neq R$ a.s. If lim sup $L_t = -\infty$ then lim inf $L_t = -\infty$ and lim $L_t = -\infty$. But this contradicts lemma 4.23b. Therefore lim sup $L_t = +\infty$ and by lemma 4.15a and $m_1 > 0$, lim $L_t = +$ a.s. and by assumption 4.12, lim $p_t = 1$ a.s.

The intuitive character of the distinction between conditions d and e of the theorem and the other three conditions should be clear. If $m_1 > 0$ then probability one of selection of strategy 1 is an absorbing barrier. If $m_1 \leq 0$ then probability one of selection of strategy 1 is a reflecting barrier. Similarly if $m_2 < 0$ then probability zero of selection of strategy 1 is an absorbing barrier. If $m_2 \geq 0$ then probability zero of selection of strategy 1 is a reflecting barrier.

Definition 4.25: A function f(x) is *strictly concave* if the tangent to f(x) at any x lies above the graph of f. That is, $f'' < 0$.

Lemma 4.26: If f is a real, continuous and strictly concave function defined on an interval (a,b) of the real line, X a real random variable defined on (a,b), then

$$E[f(X)] \leq f(E[X]).$$

Proof:

Since f is concave:

$$f(x) \leq f(E[X]) + f'(E[X])(x-E[X])$$

$$\int_a^b f(x)dP \leq \int_a^b (f(E[X]) + f'(E[X])(x-E[X]))dP$$

$$E[f(X)] < f(E[X]).$$

Corollary 4.27: The natural logarithm is a concave function and therefore $E[\ln(X)] \leq \ln(E[X])$.

Now we consider the convergence of SR1 a and b plans. Let $|S| = 2$ and suppose that strategy 1 is such that $\rho_1 = \rho^*$ (the case $\rho_2 = \rho^*$ is symmetric). Then we first need to calculate the values of m_1 and m_2 for an SR1a.

$$m_1 = E[\Delta L_{1,t}] = E[\ln\frac{\mu_{1,t}+K_1}{\rho^*+k_1}]$$

$$= E[\ln(\mu_{1,t}+k_1)] - \ln(\rho^*-k_1)$$

applying 4.27,

$$m_1 \leq 0 \ .$$

$$m_2 = E[\Delta L_{2,t}] = E[\ln\frac{\rho^*+k_1}{\mu_{2,t}+k_1}]$$

$$= \ln(\rho^*+k_1) - E[\ln(\mu_{2,t}+k_1)]$$

applying 4.27,

$$\ln(\rho^*+k_1) - \ln(\rho_2+k_1)$$

and since $\rho^* > \rho_2$

$$m_2 \geq 0 \ .$$

We now have the conditions necessary to apply theorems 4.17d, 4.24a and b and obtain:

Theorem 4.28: An SR1a plan over a pure problem basis $\langle A,\mu,S\rangle$ with $|S| = 2$ will not converge to the strategy with higher mean. In fact, $\lim\sup P_{1,t} = 1$ and $\lim\inf P_{1,t} = 0$.

Let $\tilde{\nu}_t = \max(\rho^*,\mu_{j,t})$ where j is the strategy used at time t. Then for the SR1b plan we have

$$m_2 = E[\Delta L_{2,t}] = E[\ln\frac{\tilde{\nu}_t+k_1}{\mu_{2,t}+k_1}]$$

$$= E[\ln(\tilde{\nu}_t+k_1)] - E[\ln(\mu_{2,t}+k_1)]$$

by 4.27 $\quad \geq E[\ln(\tilde{\nu}_t+k_1)] - \ln(\rho_2+k_1)$

since $\tilde{v}_t > \rho_2$, $\ln(\tilde{v}_t+k_1) > \ln(\rho_2+k_1)$ and $E[\ln(\tilde{v}_t+k_1)] > \ln(\rho_2+k_1)$

therefore,

$$m_2 \geq 0 .$$

$$m_1 = E[\ln\frac{\mu_{1,t}+k_1}{\tilde{v}_t+k_1}]$$

$$= E[\ln(\mu_{1,t}+k_1)] - E[\ln(\tilde{v}_t+k_1)]$$

$$\leq \ln(\rho^*+k_1) - E[\ln(\tilde{v}_t+k_1)]$$

since $\tilde{v}_t \geq \rho^*$, $\ln(\tilde{v}_t+k_1) \geq \ln(\rho^*+k_1)$, and $E[\ln(\tilde{v}_t+k_1)] \geq \ln(\rho^*+k_1)$

therefore,

$$m_1 \leq 0 .$$

Again applying theorem 4.17d, 4.24a and b we obtain:

Theorem 4.29: An SR1b plan over a pure problem basis $\langle A,\mu,S\rangle$ with $|S| = 2$ will not converge to the strategy with the higher mean. In fact $\limsup P_{1,t} = 1$ and $\liminf P_{1,t} = 0$.

We now investigate the convergence properties of SR3 A and B plans with $|S| = 2$. Again we assume that $\rho_1 = \rho^*$ (the case $\rho_2 = \rho^*$ is symmetric). First, we calculate the values of the $m_{i,t}$ for SR3A:

$$m_{1,t} = E[\Delta L_{1,t}] = E[\ln\frac{\lambda_{1,t}+k_1}{\rho^*+k_1}]$$

$$= E[\ln(\lambda_{1,t}+k_1)] - E[\ln(\rho^*+k_1)]$$

$$\leq \ln(E[\lambda_{1,t}+k_1]) - \ln(\rho^*+k_1)$$

by 4.27 $\qquad \leq 0 .$

For $m_{2,t}$ we find:

$$m_{2,t} = E[\Delta L_{2,t}] = E[\ln\frac{\rho^*+k_1}{\lambda_{2,t}+k_1}]$$

$$= E[\ln(\rho^*+k_1)] - E[\ln(\lambda_{2,t}+k_1)]$$

$m_{2,t} \geq \ln(\rho^*+k_1) - \ln(\rho_2+k_1)$

$m_{2,t} \geq 0$.

By similar calculations for the SR3B plans with $\theta_t = \max(\rho^*, \lambda_{j,t})$ where j is the strategy used at time t, we find that the same relations hold and $m_{1,t} \leq 0$ while $m_{2,t} \geq 0$.

Lemma 4.30: If $m_{1,t} \leq 0$, $m_{2,t} \geq 0$ for all t, $\sigma_{1,t} > 0$ and $\sigma_{2,t} > 0$ then lim sup $p_t = 1$, and lim inf $p_t = 0$.

Proof: the proof is similar to that for theorem 4.24 and relies on 4.24 part c.

Now we have shown the following:

Theorem 4.31: SR3A and SR3B plans over a pure problem basis $\langle A, \mu, S \rangle$ with $|S| = 2$ will not converge to the strategy with the higher mean. In fact lim sup $P_{1,t} = 1$ and lim inf $P_{1,t} = 0$.

CHAPTER 5

CONCLUSIONS

In this work we have examined the convergence properties of a class of probabilistic sequential adaptive schemes called sequential reproductive plans, type I (SRP). Theorem 3.21 shows that a subclass, the SR1 plan over a finite restricted deterministic arena, converges with probability 1 to a set of "good" strategies. However Theorem 3.16 shows that these plans do not converge fast enough to achieve the finite loss claimed in Holland [1970].

We have shown in Chapter 4 (Theorems 4.2, 4.3, 4.7, and 4.10) that there is not an intuitive analogue in pure problem bases for the notions of optimal strategy and arena which were defined for deterministic bases. In fact, by extending results in mathematical psychology, we have shown that several large subclasses of SRP plans do not converge in pure problem bases. (Theorems 4.28, 4.29, 4.31)

Because of the convergence problems of SRP type I plans, we suggest that they are not adequate models of the duplication process in genetic adaptation. Two further plans have been studied as a result of these findings. A plan which we call SRP Type II, has been developed by Holland and Moler [unpublished] which uses the concept of an arena as a foundation for a non probabilistic, block structured sampling scheme. This plan does overcome many of the difficulties of SRP Type I plans. However, since it relies on the concept of an arena, there are convergence problems in non-deterministic bases. Holland [1973] has examined the convergence of a much simpler implementation of the duplication process.

Extensions of this work may be made in several directions. The

results we have obtained on Linear Additive models could be extended to n dimensions. Further models of the duplication operator can be studied and incorporated in a detailed theoretical study of the other genetic operators in the artificial adaptive framework. We are presently working on a theoretical comparison of artificial genetic techniques with numerical analysis techniques in n dimensional function maximization.

APPENDIX A*

Theorem: Let X_1, X_2,... be a sequence of independent identically distributed real random variables, not identically zero. Then

$$\sum_{n=1}^{\infty} \frac{S_n}{n} \qquad \text{diverges a.s.}$$

and a certain trichotomy holds.

Proof: 1. Suppose first that $P(X_1 > 0) > 0$ and $P(X_1 < 0) > 0$. Writing $\sum_{n=1}^{m} \frac{S_n}{n} = X_1 \sum_{i=1}^{m} (1/i) + Y_m$, we note that

$$[\overline{\lim}\, Y_m < \infty,\ X_1 \sum_{t=1}^{m} (1/i) \to -\infty] \subset [\sum_{n=1}^{m} (S_n/n) \to -\infty]$$

$$[\underline{\lim}\, Y_m > \infty,\ X_1 \sum_{t=1}^{m} (1/i) \to +\infty] \subset [\sum_{n=1}^{m} (S_n/n) \to +\infty]$$

$$[\underline{\lim}\, Y_m = -\infty,\ \overline{\lim}\, Y_m = +\infty,\ X_1 \sum_{i=1}^{m} (1/i) \to -\infty] \subset [\lim \sum_{n=1}^{m} (S_n/n) \to -\infty]$$

$$[\underline{\lim}\, Y_m = -\infty,\ \overline{\lim}\, Y_m = +\infty,\ X_1 \sum_{i=1}^{m} (1/i) \to +\infty] \subset [\lim \sum_{n=1}^{m} (S_n/n) \to +\infty]$$

2. By the Hewitt-Savage 0-1 law, the four events on the right hand side of the above implications have 0 or 1 probabilities. Now $P(X_1 \sum_{i=1}^{m} (1/i) \to -\infty) > 0$ and $P(X_1 \sum_{i=1}^{m} (1/i) \to +\infty) > 0$ and at least one of the following is true: $P(\overline{\lim}\, Y_m < \infty) > 0$, $P(\underline{\lim}\, Y_m < -\infty) > 0$ or $P(\underline{\lim}\, Y_m = -\infty, \overline{\lim}\, Y_m = +\infty) > 0$.

Consequently we have the following trichotomy:

i) $\sum_{n=1}^{m} \frac{S_n}{n} \to -\infty$ a.s.;

ii) $\sum_{n=1}^{m} \frac{S_n}{n} \to +\infty$ a.s.

* See Koopmans *et al.* [to be published].

iii) $\underline{\lim}\ \Sigma \frac{S_n}{n} \to -\infty$ and $\overline{\lim}\ \Sigma \frac{S_n}{n} \to +\infty$ a.s.

Writing $\Sigma \frac{S_n}{n} = \Sigma \frac{S_n^+}{n} - \Sigma \frac{S_n^-}{n}$, we see that in case i) $\Sigma \frac{S_n^-}{n} = \infty$

in case ii) $\Sigma \frac{S_n^+}{n} \to \infty$ and in case iii) $\Sigma \frac{S_n^-}{n} = \infty$ and $\Sigma \frac{S_n^+}{n} = \infty$.

3. If $P(X_1 \leq 0) = 1$, then case i) with $\Sigma \frac{S_n}{n} = -\Sigma \frac{S_n^-}{n} = -\infty$ a.s. and $\Sigma \frac{S_n^+}{n} \equiv 0$

If $P(X_1 \quad 0) = 1$, then case ii) with $\Sigma \frac{S_n}{n} = \Sigma \frac{S_n^+}{n} = \infty$ a.s. and $\Sigma \frac{S_n^-}{n} \equiv 0$

In all three cases, $\Sigma \frac{S_n}{n}$ is divergent a.s.

Theorem:

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = 0$, and $0 < \sigma^2 < \infty$.

Let $S_n = X_1 + \ldots + X_n$, $S_n^+ = \max(0, S_n)$ and $S_n^- = \max(0, -S_n)$.

Then $\sum_1^\infty \frac{S_n^+}{n} = +\infty$ a.s. and $\sum_1^\infty \frac{S_n^-}{n} = -\infty$ a.s.

Proof: From the previous theorem it follows that either

$$\sum_1^\infty \frac{S_n^+}{n} = \infty \text{ a.s.}$$

or

$$\sum_1^\infty \frac{S_n^-}{n} = \infty \text{ a.s.}$$

Suppose that $\sum_1^\infty \frac{S_n^+}{n} = \infty$ a.s. and $\sum_1^\infty \frac{S_n^-}{n}$ does not diverge.

Then $\sum_1^\infty \frac{S_n}{n} = +\infty$ a.s.

$\therefore \quad P\left(\Sigma \frac{S_n}{n} \to +\infty\right) = 1$

Let $T_N = \sum_1^N \frac{S_n}{n} = X_1 \sum_{i=1}^N {}^1/i + \ldots + X_N \cdot {}^1/N$

$$= \sum_{k=1}^N C_{k,N} X_k$$

$$\text{where } C_{k,N} = {}^{1}/k + \ldots + {}^{1}/N$$

Consider

$$\frac{T_N}{\sqrt{\sum_{k=1}^{N} C_{k,N}^2}} = \sum_{k=1}^{N} d_{k,N} X_k = \sum_{k=1}^{N} Y_{k,N}$$

It is easy to see that

$$d_{1,N} \geq d_{2,N} \geq \ldots \geq d_{N,N}$$

and

$$\lim_{N\to\infty} d_{1,N} = 0$$

and $\therefore$ $\{Y_{k,N} \quad 1 \leq k \leq N \;, \quad N \geq 1\}$ are uniformly small random variables.

Let $F_{k,N}$ be the distribution function of $Y_{k,N}$. Then by the Lindeberg-Feller criterion:

$$\frac{T_N}{\sqrt{\sum_{k=1}^{N} C_{k,N}^2}}$$

is asymptotically normal with zero mean and unit variance,

if

$$\lim_{N\to\infty} \sum_{k=1}^{N} \int_{|y|\geq\varepsilon} y^2 dF_{k,N} = 0 \quad .$$

But

$$\sum_{k=1}^{N} \int_{|d_{k,N}X_k| \geq \varepsilon} d_{k,N} X_k^2 \; dF_k \text{ , where } F_k \text{ is the d.f. of } X_k$$

$$= \sum_{k=1}^{N} d_{k,N}^2 \int_{|d_{k,N}X_k| \geq \varepsilon} X_k^2 dF_k(X_k) \quad = \sum_{k=1}^{N} d_{k,N}^2 \int_{|d_{k,N}X| \geq \varepsilon} X^2 \; dF_1(X)$$

$$\leq \sum_{k=1}^{N} d_{k,N}^2 \int_{|d_{1,N}X| > \varepsilon} x^2 dF_1(X) \quad \leq \int_{|d_{1,N}X| > \varepsilon} X^2 dF_1(X)$$

$$\to 0 \quad \text{as } N\to\infty \qquad \text{since } \lim_{N\to\infty} d_{1,N} = 0 .$$

So the central limit theorem holds and hence

$$\lim_{N\to\infty} P[\frac{T_N}{\sqrt{\Sigma C_{k,N}{}^2}} > 0] = {}^1/2 .$$

$$\text{i.e.} \quad \lim_{N\to\infty} P[T_N > 0] = {}^1/2$$

but this contradicts what we had before, namely;

$$\lim_{N\to\infty} T_N = \sum_1^\infty \frac{S_n}{n} = +\infty \text{ a.s.}$$

Therefore, the assumption that $\Sigma \frac{S_n^-}{n}$ does not diverge is false.

Hence both

$$\Sigma \frac{S_n^-}{n} \text{ and } \Sigma \frac{S_n^+}{n} \text{ diverge a.s.}$$

REFERENCES

BEST COPY AVAILABLE

1. Bosworth, J.; Foo, N.; and Zeigler, B.P. "Comparison of Genetic Algorithms with Conjugate Gradient Methods" Computer and Communication Sciences, The University of Michigan, 1972.

2. Cavicchio, D.J. "Adaptive Search Using Simulated Evolution" Doctoral Dissertation, Computer and Communications Sciences, The University of Michigan, 1970.

3. Cover, T.M. "Hypothesis Testing with Finite Statistics" *The Annals of Mathematical Statistics, 40*, 3, pp. 828-835. 1969.

4. Foo, N.; and Bosworth, J. "Algebraic, Geometric, and Stochastic Aspects of Genetic Operators" Computer and Communication Sciences, The University of Michigan, 1972.

5. Frantz, Daniel R. "Non-linearities in Genetic Adaptive Search" Doctoral Dissertation, Computer and Communication Sciences, The University of Michigan, 1972.

6. Hellman, M.E. and Cover, T.M. "Learning with Finite Memory" *The Annals of Mathematical Statistics, 41*, 3, pp. 765-782, 1970.

7. ___, "On Memory Saved by Randomization" *The Annals of Mathematical Statistics, 42*, , pp. , 1971.

8. Holland, J. "A New Kind of Turnpike Theorem" *Bull. of the Amer. Math. Society, 75*, 6, pp. 1311-1317, 1969.

9. ___, "Robust Algorithms for Adaptation Set in a General Formal Framework" *Proceedings of the Ninth IEEE Symposium on Adaptive Processes, "Decision and Control"*, Austin, Texas, 1970.

10. ___, "Genetic Algorithms and the Optimal Allocation of Trials" *SIAM Journal, 2, 2*, 1973.

11. ___, *Introduction to Adaptation in Natural and Artificial Systems*, to be published.

12. Hollstien, R.B. "Artificial Genetic Adaptation in Computer Control Systems" Doctoral Dissertation, Computer, Information and Control Engineering, The University of Michigan, 1971.

13. Knopp, K. *Infinite Sequences and Series*, Dover Publications, New York, 1956.

14. Koopmans, L.H., N. Moler, P. Pathak, and C. Qualls, "On The Divergence of a Certain Random Series," Submitted for publication.

15. Lamperti, J. "Criteria for the Recurrence of Transience of Stochastic Process. I" *J. of Math. Analysis and Applications, 1*, pp. 314-330, 1960.

16. Lamperti, J. and Suppes, P. "Some Asymptotic Properties of Luce's Beta Learning Model" *Psychometrika*, *25*, *3*, pp. 233-241, 1960.

17. Luce, R.D. *Individual Choice Behavior*. New York: Wiley, 1959.

18. Neveu, J. *Mathematical Foundations of the Calculus of Probability*, Holden-Day, Inc. 1965.

19. Norman, M.F. "Limit Theorems for Additive Learning Models" *J. Math. Psychology*, *7*, 1, pp. 1-11, 1970.

20. Robbins, H.E. "Some Aspects of the Sequential Design of Experiments" *Bull. Am. Math. Soc.*, *55*, pp. 527-535, 1952.

21. ___, "A Sequential Decision Problem with a Finite Memory" *Proc. Nat. Acad. Sci.*, *42*, pp. 920-923, 1956.

22. Shapiro, I.J. and Narendra, K.S. "Use of Stochastic Automata for Parameter Self-Optimization with Multimodal Performance Criteria" *IEEE Trans. on SSC*, *SSC-5*, 4, pp. 352-360, 1969.

BEST COPY AVAILABLE

(Continued from inside front cover)

96 R. C. Atkinson, J. W. Brelsford, and R. M. Shiffrin. Multi-process models for memory with applications to a continuous presentation task. April 13, 1966. (J. math. Psychol., 1967, 4, 277-300).
97 P. Suppes and E. Crothers. Some remarks on stimulus-response theories of language learning. June 12, 1966.
98 R. Bjork. All-or-none subprocesses in the learning of complex sequences. (J. math. Psychol., 1968, 1, 182-195).
99 E. Gammon. The statistical determination of linguistic units. July 1, 1966.
100 P. Suppes, L. Hyman, and M. Jerman. Linear structural models for response and latency performance in arithmetic. In J. P. Hill (ed.), Minnesota Symposia on Child Psychology. Minneapolis, Minn.: 1967. Pp. 160-200).
101 J. L. Young. Effects of intervals between reinforcements and test trials in paired-associate learning. August 1, 1966.
102 H. A. Wilson. An investigation of linguistic unit size in memory processes. August 3, 1966.
103 J. T. Townsend. Choice behavior in a cued-recognition task. August 8, 1966.
104 W. H. Batchelder. A mathematical analysis of multi-level verbal learning. August 9, 1966.
105 H. A. Taylor. The observing response in a cued psychophysical task. August 10, 1966.
106 R. A. Bjork. Learning and short-term retention of paired associates in relation to specific sequences of interpresentation intervals. August 11, 1966.
107 R. C. Atkinson and R. M. Shiffrin. Some Two-process models for memory. September 30, 1966.
108 P. Suppes and C. Ihrke. Accelerated program in elementary-school mathematics--the third year. January 30, 1967.
109 P. Suppes and I. Rosenthal-Hill. Concept formation by kindergarten children in a card-sorting task. February 27, 1967.
110 R. C. Atkinson and R. M. Shiffrin. Human memory: a proposed system and its control processes. March 21, 1967.
111 Theodore S. Rodgers. Linguistic considerations in the design of the Stanford computer-based curriculum in initial reading. June 1, 1967.
112 Jack M. Knutson. Spelling drills using a computer-assisted instructional system. June 30, 1967.
113 R. C. Atkinson. Instruction in initial reading under computer control: the Stanford Project. July 14, 1967.
114 J. W. Brelsford, Jr. and R. C. Atkinson. Recall of paired-associates as a function of overt and covert rehearsal procedures. July 21, 1967.
115 J. H. Stelzer. Some results concerning subjective probability structures with semiorders. August 1, 1967
116 D. E. Rumelhart. The effects of interpresentation intervals on performance in a continuous paired-associate task. August 11, 1967.
117 E. J. Fishman, L. Keller, and R. E. Atkinson. Massed vs. distributed practice in computerized spelling drills. August 18, 1967.
118 G. J. Groen. An investigation of some counting algorithms for simple addition problems. August 21, 1967.
119 H. A. Wilson and R. C. Atkinson. Computer-based instruction in initial reading: a progress report on the Stanford Project. August 25, 1967.
120 F. S. Roberts and P. Suppes. Some problems in the geometry of visual perception. August 31, 1967. (Synthese, 1967, 17, 173-201)
121 D. Jamison. Bayesian decisions under total and partial ignorance. D. Jamison and J. Kozielecki. Subjective probabilities under total uncertainty. September 4, 1967.
122 R. C. Atkinson. Computerized instruction and the learning process. September 15, 1967.
123 W. K. Estes. Outline of a theory of punishment. October 1, 1967.
124 T. S. Rodgers. Measuring vocabulary difficulty: An analysis of item variables in learning Russian-English and Japanese-English vocabulary parts. December 18, 1967.
125 W. K. Estes. Reinforcement in human learning. December 20, 1967.
126 G. L. Wolford, D. L. Wessel, W. K. Estes. Further evidence concerning scanning and sampling assumptions of visual detection models. January 31, 1968.
127 R. C. Atkinson and R. M. Shiffrin. Some speculations on storage and retrieval processes in long-term memory. February 2, 1968.
128 John Holmgren. Visual detection with imperfect recognition. March 29, 1968.
129 Lucille B. Mlodnosky. The Frostig and the Bender Gestalt as predictors of reading achievement. April 12, 1968.
130 P. Suppes. Some theoretical models for mathematics learning. April 15, 1968 (Journal of Research and Development in Education, 1967, 1, 5-22)
131 G. M. Olson. Learning and retention in a continuous recognition task. May 15, 1968.
132 Ruth Norene Hartley. An investigation of list types and cues to facilitate initial reading vocabulary acquisition. May 29, 1968.
133 P. Suppes. Stimulus-response theory of finite automata. June 19, 1968.
134 N. Moler and P. Suppes. Quantifier-free axioms for constructive plane geometry. June 20, 1968. (In J. C. H. Gerretsen and F. Oort (Eds.), Compositio Mathematica. Vol. 20. Groningen, The Netherlands: Wolters-Noordhoff, 1968. Pp. 143-152.)
135 W. K. Estes and D. P. Horst. Latency as a function of number of response alternatives in paired-associate learning. July 1, 1968.
136 M. Schlag-Rey and P. Suppes. High-order dimensions in concept identification. July 2, 1968. (Psychom. Sci., 1968, 11, 141-142)
137 R. M. Shiffrin. Search and retrieval processes in long-term memory. August 15, 1968.
138 R. D. Freund, G. R. Loftus, and R.C. Atkinson. Applications of multiprocess models for memory to continuous recognition tasks. December 18, 1968.
139 R. C. Atkinson. Information delay in human learning. December 18, 1968.
140 R. C. Atkinson, J. E. Holmgren, and J. F. Juola. Processing time as influenced by the number of elements in the visual display. March 14, 1969.
141 P. Suppes, E. F. Loftus, and M. Jerman. Problem-solving on a computer-based teletype. March 25, 1969.
142 P. Suppes and Mona Morningstar. Evaluation of three computer-assisted instruction programs. May 2, 1969.
143 P. Suppes. On the problems of using mathematics in the development of the social sciences. May 12, 1969.
144 Z. Domotor. Probabilistic relational structures and their applications. May 14, 1969.
145 R. C. Atkinson and T. D. Wickens. Human memory and the concept of reinforcement. May 20, 1969.
146 R. J. Titiev. Some model-theoretic results in measurement theory. May 22, 1969.
147 P. Suppes. Measurement: Problems of theory and application. June 12, 1969.
148 P. Suppes and C. Ihrke. Accelerated program in elementary-school mathematics--the fourth year. August 7, 1969.
149 D. Rundus and R.C. Atkinson. Rehearsal in free recall: A procedure for direct observation. August 12, 1969.
150 P. Suppes and S. Feldman. Young children's comprehension of logical connectives. October 15, 1969.

(Continued on back cover)

BEST COPY AVAILABLE

(Continued from inside back cover)

151 Joaquim H. Laubsch. An adaptive teaching system for optimal item allocation. November 14, 1969.
152 Roberta L. Klatzky and Richard C. Atkinson. Memory scans based on alternative test stimulus representations. November 25, 1969.
153 John E. Holmgren. Response latency as an indicant of information processing in visual search tasks. March 16, 1970.
154 Patrick Suppes. Probabilistic grammars for natural languages. May 15, 1970.
155 E. Gammon. A syntactical analysis of some first-grade readers. June 22, 1970.
156 Kenneth N. Wexler. An automaton analysis of the learning of a miniature system of Japanese. July 24, 1970.
157 R. C. Atkinson and J. A. Paulson. An approach to the psychology of instruction. August 14, 1970.
158 R.C. Atkinson, J.D. Fletcher, H.C. Chetin, and C.M. Stauffer. Instruction in initial reading under computer control: the Stanford project. August 13, 1970.
159 Dewey J. Rundus. An analysis of rehearsal processes in free recall. August 21, 1970.
160 R.L. Klatzky, J.F. Juola, and R.C. Atkinson. Test stimulus representation and experimental context effects in memory scanning.
161 William A. Rottmayer. A formal theory of perception. November 13, 1970.
162 Elizabeth Jane Fishman Loftus. An analysis of the structural variables that determine problem-solving difficulty on a computer-based teletype. December 18, 1970.
163 Joseph A. Van Campen. Towards the automatic generation of programmed foreign-language instructional materials. January 11, 1971.
164 Jamesine Friend and R.C. Atkinson. Computer-assisted instruction in programming: AID. January 25, 1971.
165 Lawrence James Hubert. A formal model for the perceptual processing of geometric configurations. February 19, 1971.
166 J. F. Juola, I.S. Fischler, C.T. Wood, and R.C. Atkinson. Recognition time for information stored in long-term memory.
167 R.L. Klatzky and R.C. Atkinson. Specialization of the cerebral hemispheres in scanning for information in short-term memory.
168 J.D. Fletcher and R.C. Atkinson. An evaluation of the Stanford CAI program in initial reading (grades K through 3). March 12, 1971.
169 James F. Juola and R.C. Atkinson. Memory scanning for words versus categories.
170 Ira S. Fischler and James F. Juola. Effects of repeated tests on recognition time for information in long-term memory.
171 Patrick Suppes. Semantics of context-free fragments of natural languages. March 30, 1971.
172 Jamesine Friend. Instruct coders' manual. May 1, 1971.
173 R.C. Atkinson and R.M. Shiffrin. The control processes of short-term memory. April 19, 1971.
174 Patrick Suppes. Computer-assisted instruction at Stanford. May 19, 1971.
175 D. Jamison, J.D. Fletcher, P. Suppes, and R.C.Atkinson. Cost and performance of computer-assisted instruction for compensatory education.
176 Joseph Offir. Some mathematical models of individual differences in learning and performance. June 28, 1971.
177 Richard C. Atkinson and James F. Juola. Factors influencing speed and accuracy of word recognition. August 12, 1971.
178 P. Suppes, A. Goldberg, G. Kanz, B. Searle, and C. Stauffer. Teacher's handbook for CAI courses. September 1, 1971.
179 Adele Goldberg. A generalized instructional system for elementary mathematical logic. October 11, 1971.
180 Max Jerman. Instruction in problem solving and an analysis of structural variables that contribute to problem-solving difficulty. November 12, 1971.
181 Patrick Suppes. On the grammar and model-theoretic semantics of children's noun phrases. November 29, 1971.
182 Georg Kreisel. Five notes on the application of proof theory to computer science. December 10, 1971.
183 James Michael Moloney. An investigation of college student performance on a logic curriculum in a computer-assisted instruction setting. January 28, 1972.
184 J.E. Friend, J.D. Fletcher, and R.C. Atkinson. Student performance in computer-assisted instruction in programming. May 10, 1972.
185 Robert Lawrence Smith, Jr. The syntax and semantics of ERICA. June 14, 1972.
186 Adele Goldberg and Patrick Suppes. A computer-assisted instruction program for exercises on finding axioms. June 23, 1972.
187 Richard C. Atkinson. Ingredients for a theory of instruction. June 26, 1972.
188 John D. Bonvillian and Veda R. Charrow. Psycholinguistic implications of deafness: A review. July 14, 1972.
189 Phipps Arabie and Scott A. Boorman. Multidimensional scaling of measures of distance between partitions. July 26, 1972.
190 John Ball and Dean Jamison. Computer-assisted instruction for dispersed populations: System cost models. September 15, 1972.
191 W. R. Sanders and J. R. Ball. Logic documentation standard for the Institute for Mathematical Studies in the Social Sciences. October 4, 1972.
192 M.T. Kane. Variability in the proof behavior of college students in a CAI course in logic as a function of problem characteristics. October 6, 1972.
193 P. Suppes. Facts and fantasies of education. October 18, 1972.
194 R. C. Atkinson and J. F. Juola. Search and decision processes in recognition memory. October 27, 1972.
195 P. Suppes, R. Smith, and M. Léveillé. The French syntax and semantics of PHILIPPE, part 1 : Noun phrases. November 3, 1972.
196 D. Jamison, P. Suppes, and S. Wells. The effectiveness of alternative instructional methods: A survey. November 1972.
197 P. Suppes. A survey of cognition in handicapped children. December 29, 1972.
198 B. Searle, P. Lorton, Jr., A. Goldberg, P. Suppes, N. Ledet, and C. Jones. Computer-assisted instruction program: Tennessee State University. February 14, 1973.
199 D. R. Levine. Computer-based analytic grading for German grammar instruction. March 16, 1973.
200 P. Suppes, J.D. Fletcher, M. Zanotti, P. V. Lorton, Jr., and B. W. Searle. Evaluation of computer-assisted instruction in elementary mathematics for hearing-impaired students. March 17, 1973.
201 G. A. Huff. Geometry and formal linguistics. April 27, 1973.
202 C. Jensema. Useful techniques for applying latent trait mental-test theory. May 9, 1973.
203 A. Goldberg. Computer-assisted instruction. The application of theorem-proving to adaptive response analysis. May 25, 1973.
204 R. C. Atkinson, D. J. Herrmann, and K. T. Wescourt. Search processes in recognition memory. June 8, 1973.
205 J. Van Campen. A computer-based introduction to the morphology of Old Church Slavonic. June 18, 1973.
206 R. B. Kimball. Self-optimizing computer-assisted tutoring: Theory and practice. June 25, 1973.
207 R.C. Atkinson, J.D. Fletcher, E. J. Lindsay, J.O. Campbell, and A. Barr. Computer-assisted instruction in initial reading. July 9, 1973.
208 V. R. Charrow and J. D. Fletcher. English as the second language of deaf students. July 20, 1973.
209 J. A. Paulson. An evaluation of instructional strategies in a simple learning situation. July 30, 1973.
210 N. Martin. Convergence properties of a class of probabilistic adaptive schemes called sequential reproductive plans. July 31, 1973.